Article

Quantitative Evaluation of CFTR Pre-mRNA Splicing Dependent on the (TG)mTn Poly-Variant Tract

Manuela Sterrantino [1], **Andrea Fuso** [1], **Silvia Pierandrei** [1], **Sabina Maria Bruno** [1], **Giancarlo Testino** [1], **Giuseppe Cimino** [2], **Antonio Angeloni** [1] **and Marco Lucarelli** [1,3,*]

1   Department of Experimental Medicine, Sapienza University of Rome, 00161 Rome, Italy; manusterr@hotmail.it (M.S.); andrea.fuso@uniroma1.it (A.F.); pierandrei.silvia@gmail.com (S.P.); sabinamaria.bruno@gmail.com (S.M.B.); giancarlo.testino@uniroma1.it (G.T.); antonio.angeloni@uniroma1.it (A.A.)
2   Cystic Fibrosis Reference Center of Lazio Region, AOU Policlinico Umberto I, 00161 Rome, Italy; ciminolo@tiscali.it
3   Pasteur Institute, Cenci Bolognetti Foundation, Sapienza University of Rome, 00161 Rome, Italy
*   Correspondence: marco.lucarelli@uniroma1.it

Abstract: Genetic analysis in cystic fibrosis (CF) is a difficult task. Within the many causes of variability and uncertainty, a major determinant is poor knowledge of the functional effect of most DNA variants of the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) gene. In turn, knowledge of the effect of a CFTR variant has dramatic diagnostic, prognostic and, in the era of CF precision medicine, also therapeutic consequences. One of the most challenging CFTR variants is the (TG)mTn haplotype, which has variable functional effect and controversial clinical consequences. The exact quantification of the anomalous splicing of CFTR exon 10 (in the HGVS name; exon 9 in the legacy name) and, consequently, of the residual wild-type functional CFTR mRNA, should be mandatory in clinical assessment of patients with potentially pathological haplotype of this tract. Here, we present a real time-based assay for the quantification of the proportion of exon 10+/exon 10− CFTR mRNA, starting from nasal brushing. Our assay proved rapid, economic and easy to perform. Specific primers used for this assay are either disclosed or commercially available, allowing any laboratory to easily perform it. A simplified analysis of the data is provided, facilitating the interpretation of the results. This method helps to enhance the comprehension of the genotype–phenotype relationship in CF and CFTR-related disorders (CFTR-RD), crucial for the diagnosis, prognosis and personalized therapy of CF.

Keywords: Cystic Fibrosis; Cystic Fibrosis Transmembrane conductance Regulator (CFTR); CFTR-related disorders (CFTR-RD); (TG)mTn tract; functional effect of CFTR variants; pre-mRNA splicing; real time; nasal brushing

1. Introduction

Mutational search in the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) gene and subsequent data analysis have been greatly enhanced [1–7]. Nevertheless, the correct scheduling of genetic analysis in cystic fibrosis (CF) remains a difficult task [8]. Similarly, although a great enhancement in the approaches for CFTR functional studies has been achieved [9], the characterization from a clinical point of view of variants found during extensive genetic searches in the CFTR gene often remains uncertain. Both the organization of mutational search and the reliability of functional characterization have a great impact on the comprehension of the genotype–phenotype relationship and on the decisional process in CF [10,11]. The uncertainty and variability of the functional effect of a CFTR gene variant may reflect on the CF clinical expression [12,13] and may impair neonatal screening [14], diagnostic assessment and carrier search [8,15–17], as well as prognostic ability [18]. With the recent advent of precision medicine in CF, a full assessment of the

molecular mechanism of a CFTR pathogenic variant became mandatory to select the correct modulatory therapy [19,20]. The (TG)mTn tract in the CFTR gene has variable clinical effects, from mono-symptomatic to oligo- and poly-symptomatic [21–24]. The (TG)mTn tract has also been found to be involved in highly disruptive rearrangements [25] as well as in complex alleles [9,14,22,26–31]. Depending on the length of (TG)m and Tn repeats, different amounts of exclusion of exon 10 (in the HGVS name; exon 9 in the legacy name) from the pre-mRNA of the CFTR gene occur [32,33]. In general, the combination of a long (TG)m and a short Tn (for example, a (TG)13T5 allele) is associated with a larger amount of anomalous splicing without exon 10, whereas a short (TG)m combined with a long Tn (for example a (TG)10T9 allele) is associated with a larger amount of wild-type CFTR mRNA correctly spliced. However, it is difficult to predict the exact quantitative splicing effect of a specific (TG)mTn haplotype. For a correct functional characterization of the (TG)mTn tract, the experimental evaluation of the amount of CFTR exon 10 anomalously spliced is mandatory. The measurement of the exact ratio between correctly and incorrectly spliced CFTR mRNA depending on the (TG)mTn haplotype, allows the assessment of the residual functionality of CFTR that, in turn, is linked to clinical manifestations. A high exon 10+ to exon 10− ratio has no pathological effect, whereas a low or very low exon 10+ to exon 10− ratio usually has pathological effects spanning from, respectively, CFTR-related disorders (CFTR-RD) to CF. In this paper we present a rapid, economic and easy to perform assay, based on real time PCR and with a simplified data analysis, able to quantify the proportion of CFTR mRNA with or without exon 10.

2. Materials and Methods

Clinical samples used in this study were obtained during institutional diagnostic procedures; investigation described here could be carried out on residual specimens following diagnostic analysis with all data kept anonymous.

The CFTR genotype and (TG)mTn tracts were characterized by cycle sequencing on an ABI PRISM 3130*xl* genetic analyzer (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) as previously described [34,35]. Genetic characterization was completed by multiplex ligation-dependent probe amplification (SALSA MLPA probemix P091 CFTR, MRC Holland, Amsterdam, The Netherlands).

RNA was obtained from the nasal epithelium of CF patients, through nasal brushing. It was extracted by the RNeasy mini kit (Qiagen, Hilden, Germany) and reverse transcribed by the iScript cDNA Synthesis kit (Bio-Rad, Hercules, CA, USA) that includes a mix of oligo(dT) and random hexamers as a priming strategy. Retrotranscription was performed using 1 μg of total RNA in 10 μL, 4 μL of 5× iScript reaction mix, 1 U of iScript reverse transcriptase in 1 μL, 5 μL of H_2O, in a final volume of 20 μL, according to the manufacturer's instructions. The reactions were incubated in a PTC 100 thermocycler (Bio-Rad), according to a program that allows the synthesis of the double stranded cDNA: 5′ 25 °C, 30′ 42 °C and 5′ 85 °C.

The real time polymerase chain reaction (PCR) was carried out using two TaqMan probes and their corresponding primers, by a specific no-ROX Master Mix (FluoCycle™ II Master Mix for probe, EuroClone, Pero (MI), Italy) according to the manufacturer's instructions. The TaqMan probes are FAM dye-labeled and are suitable to study the quantity of CFTR mRNA with or without exon 10 (in the HGVS name; exon 9 in the legacy name). The exon 10+ probe spans the junction between exon 9 and exon 10 and works when exon 10 is present in the cDNA; the exon 10− probe spans the junction between exon 9 and exon 11 and works when exon 10 is absent in the cDNA. The TaqMan assay specific for exon 10+ cDNA was already commercially available (TaqMan gene expression assay, code 4448892, ID Hs01565545-m1; ThermoFisher Scientific, Waltham, MA, USA). The TaqMan assay specific for exon 10− cDNA underwent the following customized design: forward primer 5′-GTAGTGATGGAGAATGTAACAGCCT-3′ (900 nM final concentration), reverse primer 5′-GCTCCAGTTCTCCCATAATCA**Y**CAT-3′ (900 nM final concentration), FAM probe 5′-CTGGGAGGAGACTTCA-3′ (250 nM final concentration) (all synthesized

by TermoFisher Scientific, Waltham, MA, USA). The final reaction volume was 20 μL, using 1 μL of cDNA mix, 10 μL of 2× no-ROX master mix, 1 μL of specific TaqMan probe assay, 8 μL of H_2O, according to manufacturer instructions. The real time PCR instrument used was the CFX Connect (Bio-Rad), with the following program: 5′ 95 °C and 45 cycles of 15″ 95 °C followed by 1′ 60 °C. The threshold cycles (Ct) of both wild-type (exon 10+) and mutated (exon 10−) assays were acquired, in triplicate for each sample, the analysis was performed using the ΔCt, calculated as the difference between the average Ct of the exon 10− assay and the average Ct of the exon 10+ assay. Melting curves at the end of each protocol assured the specificity of both the exon 10− and exon 10+ assays.

For this work we employed an absolute reference, through the use of plasmids pCR 2.1 (Invitrogen, Waltham, MA, USA), containing exon 10 (PL 10+) or not containing it (PL 10−). For the construction of the plasmids, the CFTR zone from exon 9 to exon 11 was amplified as described below for the endpoint PCR assay. The amplicons, with or without exon 10, were cloned into the pCR 2.1 (Invitrogen, Waltham, MA, USA) by the TA cloning system (Invitrogen, Waltham, MA, USA) according to the manufacturer's instruction. The correct sequences (with or without exon 10) were verified by cycle sequencing (ThermoFisher Scientific, Waltham, MA, USA). Therefore, we created plasmid mixes with known concentrations of both PL 10+ and PL 10− and used them to construct a calibration curve. All samples were analyzed using this calibration curve. The calculation needed for the assessment of a relative quantity of exon 10+ and exon 10− CFTR mRNA, based on the calibration curve and taking into account the efficiency of exon 10+ and exon 10− assays, are described in the results section.

After the setup of the method, the analysis of 13 subjects, selected according to their (TG)mTn haplotypes and regardless of their CFTR genotype or clinical condition, was performed (Table 1). For clarity of the result, we decided to investigate only homozygous (TG)mTn subjects to study the contribution of each (TG)mTn haplotype to the final proportion of exon 10+ and exon 10−.

Besides the real time PCR assay, the expression was also assessed through an endpoint PCR assay based on agarose gel electrophoresis and semi-quantitative densitometric analysis. For the reverse transcriptase step, the same conditions described above were used. For the PCR step, the portion of cDNA containing exon 10 was amplified by using a primer located in exon 9 (forward primer: 5′-ACAAAAGCAAGAATATAAGACATTG-3′) paired with a primer located in exon 11 (reverse primer: 5′-GAATGAAATTCTTCCACTGTGC-3′). The PCR mix was in a final volume of 15 μL containing: 2.5 μL of cDNA, 175 μM of each dNTP (Fermentas, Waltham, MA, USA), 1.5 mM $MgCl_2$, 6 pmol of each primer and 0.5 U GoTaq hot start polymerase with 1X manufacturer's buffer (Promega, Madison, Wisconsin, USA). The PCR step was conducted in a PTC100 thermocycler (BIO-RAD, Hercules, CA, USA) with the following PCR cycle: 2′ 95 °C; 35 cycles of 45″ 94 °C, 1′ 30″ 60 °C, 2′ 30″ 72 °C followed by 7′ 72 °C. The endpoint PCR assay products were subsequently analyzed by electrophoresis on a 1.5% agarose gel, to detect the possible alternative splicing forms. Both amplicons, with or without exon 10, were recovered from agarose gel and their identity verified by cycle sequencing (ThermoFisher Scientific, Waltham, MA, USA). The corresponding bands were scanned by a CCD camera (VisiDoc-It; UVP, Upland, CA, USA) and examined on the bioimaging and analysis system VisionWorks LS (software version 6.7.3; UVP) for densitometric assays.

Table 1. Genotypes and splicing proportion.

ID	Diagnosis	Genotype (Legacy Name)	(TG)mTn (Legacy Name)	Exon 10 Splicing Percentages			
				Individual Percentages		Average Percentages ± ds	
				Exon 10+	Exon 10−	Exon 10+	Exon 10−
1	CF-PS	F508del/1585-9412A>G *	(TG)10T9/(TG)10T9	97.8	2.2	94.7 ± 2.7	5.3 ± 2.7
2	CF-PI	F508del/F508del	(TG)10T9/(TG)10T9	93.5	6.5		
3	CF-PI	[F508del;I1027T]/F508del	(TG)10T9/(TG)10T9	92.9	7.1		
4	Uncertain	G576A/G576A	(TG)10T7/(TG)10T7	91.3	7.7	88.0 ± 4.6	12.0 ± 4.6
5	Healthy (carrier)	CFTRdup19/+	(TG)10T7/(TG)10T7	84.7	15.3		
6	CF-PI	R553X/CFTRdele2	(TG)11T7/(TG)11T7	83.1	16.9	80.9 ± 2.2	19.1 ± 2.2
7	CFTR-RD	unknown/unknown	(TG)11T7/(TG)11T7	81.4	18.6		
8	Healthy (carrier)	R553X/+	(TG)11T7/(TG)11T7	79.6	20.4		
9	Healthy (carrier)	G85E/+	(TG)11T7/(TG)11T7	77.7	22.3		
10	Healthy (carrier)	G85E/+	(TG)11T7/(TG)11T7	82.4	17.6		
11	Healthy (gen pop)	+/+	(TG)11T5/(TG)11T5	47.5	52.5	47.5 ± 0.0	52.5 ± 0.0
12	CFTR-RD	359insT/+	(TG)12T5/(TG)12T5	36.2	63.8	27.6 ± 12.1	72.4 ± 12.1
13	CFTR-RD	+/+	(TG)12T5/(TG)12T5	19.1	80.9		

The characteristics of the 13 homozygous (TG)mTn subjects analyzed are shown. In the first and second column the identification number (ID) and the diagnosis are reported, respectively. The genotypes (third and fourth columns) are shown, as well as the individual and average percentages of exon 10+ (wild-type) and exon 10− (mutated) CFTR mRNA (columns from fifth to eighth). Only the legacy name of CFTR variants is reported in the table. The HGVS name are as follows: F508del (legacy name), c.1521_1523delCTT (HGVS cDNA name), p.Phe508del (HGVS protein name); c.1585-9412A>G (* only in HGVS cDNA name); I1027T (legacy name), c.3080T>C (HGVS cDNA name), p.Ile1027Thr (HGVS protein name); G576A (legacy name), c.1727G>C (HGVS cDNA name), p.Gly576Ala (HGVS protein name); CFTRdup19 (legacy name), c.3717+5032_3717+5033insCAins3469-52_3717+5032 (HGVS cDNA name); R553X (legacy name), c.1657C>T (HGVS cDNA name), p.Arg553* (HGVS protein name); CFTRdele2 (legacy name), c.54-1161_164+1603del2875 (HGVS cDNA name); G85E (legacy name), c.254G>A (HGVS cDNA name), p.Gly85Glu (HGVS protein name); 359insT (legacy name), c.227_228insT (HGVS cDNA name), p.Trp79Leufs*32 (HGVS protein name); (TG)10T9 (legacy name), c.[1210-34TG[10];1210-12T[9]] (HGVS cDNA name); (TG)10T7 (legacy name), c.[1210-34TG[10];1210-12T[7]] (HGVS cDNA name); (TG)11T7 (legacy name), c.[1210-34TG[11];1210-12T[7]] (HGVS cDNA name); (TG)11T5 (legacy name), c.[1210-34TG[11];1210-12T[5]] (HGVS cDNA name); (TG)12T5 (legacy name), c.[1210-34TG[12];1210-12T[5]] (HGVS cDNA name). Abbreviations: CF-PS = cystic fibrosis with pancreas sufficiency; CF-PI = cystic fibrosis with pancreas insufficiency; CFTR-RD = cystic fibrosis-related disorders; gen. pop. = general population; + = wild type. Uncertain = uncertain diagnosis; unknown/unknown = no CFTR pathological variants found.

3. Results

As a first step, we assessed the efficiency of exon 10+ and exon 10− real-time PCR assays. The efficiency of the assay with the probe 10+ was calculated using different concentrations of the plasmid carrying the exon 10 (PL 10+) with dilutions at 1 ng/μL, 0.5 ng/μL, 0.25 ng/μL and 0.125 ng/μL. The curve obtained plotting the logarithmic quantity respect to the Ct showed a value of 99% for this probe (Figure 1A). The efficiency of the assay with the probe 10− was calculated in the same way, using the different concentrations of plasmid without exon 10 (PL 10−) obtaining an efficiency of 119% (Figure 1B). The high efficiencies of the assays with the two different probes highlighted the suitability of this method to quantitatively investigate the splicing levels depending on the (TG)mTn haplotype of the CFTR gene. However, as the efficiencies of exon 10+ and exon 10− assays cannot be considered equivalent, their values will be part of the calculation.
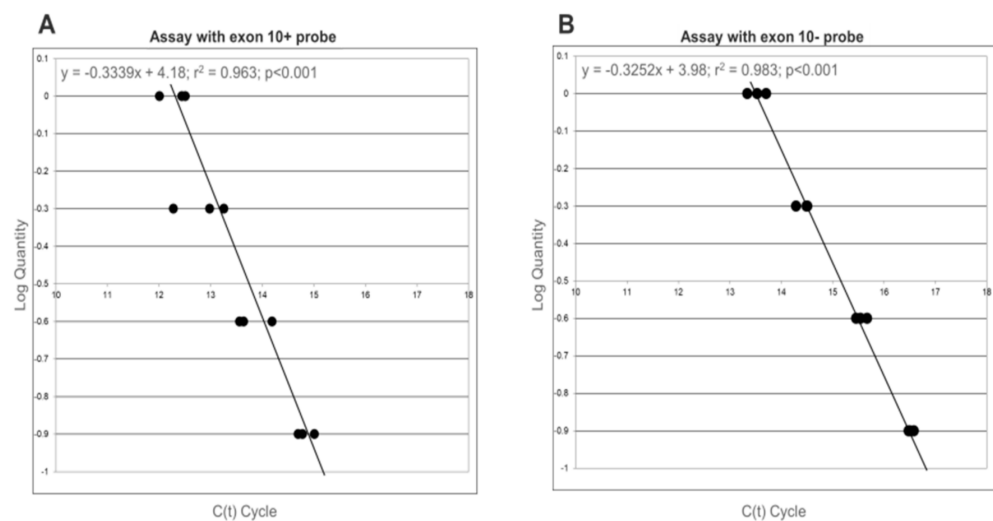

Figure 1. Efficiency curve for exon 10+ and exon 10− plasmid assays. (**A**) The efficiency of probe at junction exon 9—exon 10 is calculated using different concentrations of plasmid with exon 10 (assay with 10+ probe). (**B**) The efficiency of probe at junction exon 9—exon 11 is calculated using different concentrations of plasmid without exon 10 (assay with 10− probe). For both, the efficiency curve is obtained evaluating the logarithmic quantity in respect to the Ct cycle.

The further step has been the construction of the reference plasmid curve (Figure 2), using mixtures of plasmids, at known copy number, containing exon 10 (PL 10+) and not containing exon 10 (PL 10−). The curve was based on the logarithmic PL 10+/PL 10− plasmid ratio. We obtained two different regression lines with two different slopes, depending on the percentage of exon 10−. In particular, it appeared suitable using a regression line for samples with exon 10− splicing between 1% and 10% (the leftmost line in Figure 2) and a different regression line for samples with exon 10− splicing between 10% and 99% (the rightmost curve in Figure 2). The equations of the regression lines are indicated in Figure 2. These can be considered the calibration curves of this assay, which showed a dynamic range from 1% to 99% of exon 10 exclusion (consequently, from 99% to 1% of exon 10 inclusion). Moreover, we evaluated 10 controls with known percentages of PL 10− and PL 10+ (exactly with 2.5%, 4.0%, 7.5%, 9,0%, 25.0%, 26.0%, 50.0%, 55.0%, 80.0% and 85.0% of exon 10−), in four different experiments, to further verify the reliability of the assay (the circles in Figure 2). The formulas that can be used for the calculation of exon 10+/exon 10− proportion, according to regressions reported in Figure 2, are indicated in Table 2. An excel tool for a simplified calculation is linked to this paper as Supplementary Material.
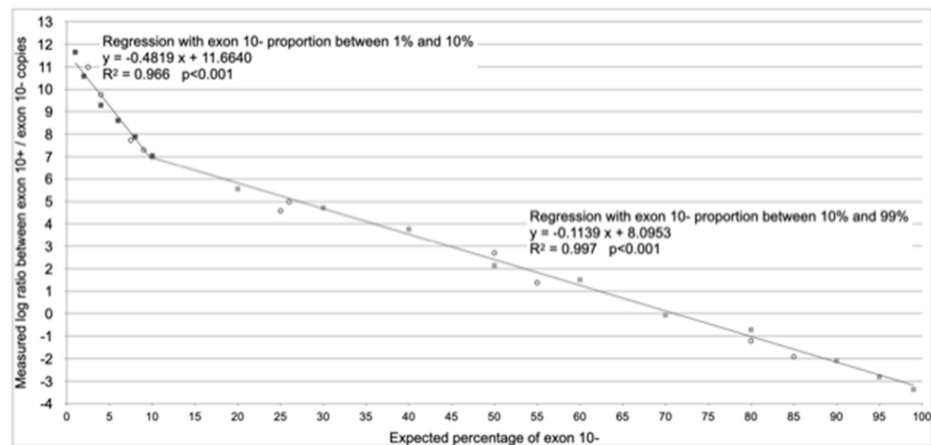
Figure 2. Reference plasmid curves and validation controls. Plasmidic reference curves of the measured logarithmic relationship between known proportions of plasmids containing exon 10 and plasmids not containing exon 10 (full squares) are shown. Also, validation controls with known ratio of exon 10+ and exon 10− (empty circles) are shown, used to confirm the reliability of the method based on the reference curves. The leftmost curve (black squares) is used to estimate the percentages of splicing without exon 10 between 1% and 10%, with the formula y = −0.4819 x + 11.6640, while the rightmost curve (grey squares) is used to estimate the percentages of splicing without exon 10 between 10% and 99%, with the formula y = −0.1139 x + 8.0953.

Table 2. Formulas for the calculation of the percentages of splicing without exon 10.

Range of Applicability	From 1% to 10% of Splicing without Exon 10 (Exon 10− Assay)	From 10% to 99% of Splicing without Exon 10 (Exon 10− Assay)
Regression equation (according to Figure 2)	y = −0.4819 x + 11.6640	y = −0.1139 x + 8.0953
Reverse regression equation (x = % of splicing without exon 10)	x = (y − 11.6640)/−0.4819	x = (y − 8.0953)/−0.1139
Formula with generic efficiency	% splicing without exon 10 = $\dfrac{\text{Log}_2\left(\dfrac{(1+\text{E10}-)^{\text{Ct10}-}}{(1+\text{E10}+)^{\text{Ct10}+}}\right) - 11.6640}{-0.4819}$	% splicing without exon 10 = $\dfrac{\text{Log}_2\left(\dfrac{(1+\text{E10}-)^{\text{Ct10}-}}{(1+\text{E10}+)^{\text{Ct10}+}}\right) - 8.0953}{-0.1139}$
Formula with experimental values of efficiencies	% splicing without exon 10 = $\dfrac{\text{Log}_2\left(\dfrac{(2.1886)^{\text{Ct10}-}}{(1.9884)^{\text{Ct10}+}}\right) - 11.6640}{-0.4819}$	% splicing without exon 10 = $\dfrac{\text{Log}_2\left(\dfrac{(2.1886)^{\text{Ct10}-}}{(1.9884)^{\text{Ct10}+}}\right) - 8.0953}{-0.1139}$

Depending on the proportion of exon 10 skipping, the different formulas indicated should be utilized. The different efficiencies of exon 10+ and exon 10− assay should be considered. Both the general formulas including generic real-time efficiencies (E10− = efficiency of the exon 10− assay; E10+ = efficiency of the exon 10+ assay), as well as specific formulas utilizing the efficiencies experimentally measured in our assays (E10− = 1.1886; E10+ = 0.9884) are reported. Ct10− = threshold cycle of the exon 10− assay; Ct10+ = threshold cycle of the exon 10+ assay.

This new assay allowed us to characterize the individual and average percentages of exon 10 splicing in 13 homozygous (TG)mTn subjects, as reported in Table 1. They were classified according to their (TG)mTn genotype and ordered in the table according to increasing average exon 10− splicing values. In Table 1 is also indicated the final diagnosis. The quantification of the final percentages of wild-type (exon 10+) and mutated (exon 10−) mRNA for each (TG)mTn genotype under examination was performed referring to a total contribution of 100%. As reported in Figure 3 and in Table 1, the highest average percentage of wild-type mRNA was shown by the (TG)10T9/(TG)10T9 genotype, with a 94.7% of exon 10+ and a 5.3% of exon 10−, while the average lowest percentage of wild-type mRNA was shown by the (TG)12T5/(TG)12T5 genotype, with a 27.6% of exon 10+ and a 72.4% of exon 10−. For the other genotypes, the average percentages were as follows: 88.0% of exon 10+ and 12.0% of exon 10− for the (TG)10T7/(TG)10T7 genotype; 80.9% of exon 10+ and 19.1%

of exon 10− for the (TG)11T7/(TG)11T7 genotype; 47.5% of exon 10+ and 52.5% of exon 10− for the (TG)11T5/(TG)11T5 genotype.
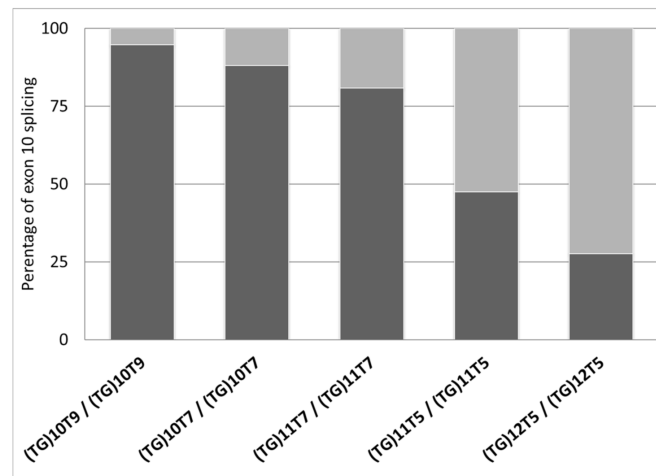

Figure 3. Relative splicing quantification of (TG)mTn genotypes. The average percentage of both wild-type (exon 10+; dark grey part of columns) and mutated (exon 10−; light grey part of columns) CFTR mRNA in the homozygous (TG)mTn genotypes analyzed is shown; the overall contribution of each genotype is set to 100%.

The results obtained through real time PCR analysis were compared with those obtained by the endpoint PCR assay, agarose gel electrophoresis and semi-quantitative densitometric analysis. The correlation between these two approaches in the quantification of the splicing amounts depending on the (TG)mTn haplotypes, for both exon 10+ and exon 10− CFTR mRNA, is reported in Figure 4. An excellent correlation (R^2 = 0.989, *p* < 0.01) was evidenced for the study of the (TG)mTn haplotype-dependent splicing by the two approaches. It should be taken under consideration that the experimental measurement of exon 10+ and exon 10− splicing is performed by two independent assays (respectively, with exon 10+ or exon 10− probe), but that the calculation of the splicing proportion is obtained by a unique formula including the resulting Ct ratio from both assays. Consequently, the two regression lines shown in Figure 4 are constructed starting from one data set (for example the percentage of exon 10− mRNA) being the other data set the complement to 100%. For this reason only one value of R^2 and p is shown. Due to complementarity, the range of exon 10+ and exon 10− mRNA are obviously different and, for clarity, both regression lines are shown in the figure.
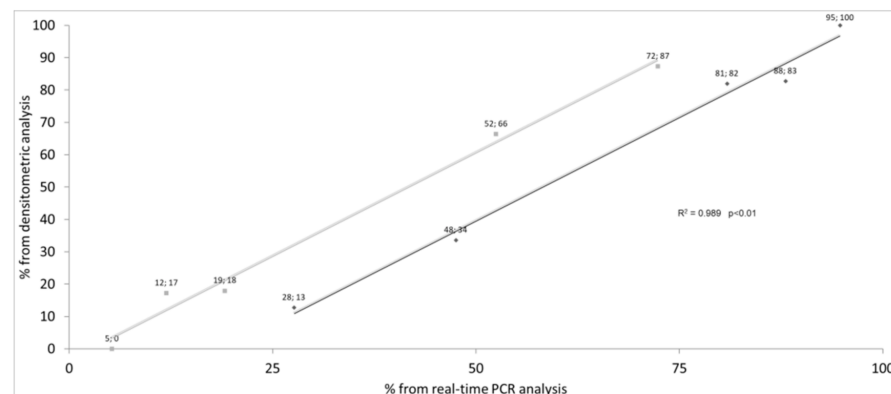

Figure 4. Correlation between densitometric and real time polymerase chain reaction (PCR) analysis. Each (TG)mTn genotype is represented as a point, both for exon 10+ (dark grey) and for exon 10− (light grey). On the x and y axes, the average percentage values of exon 10+ and exon 10− obtained by, respectively, real time PCR and densitometric analysis are reported.

4. Discussion

The (TG)mTn is a CFTR polyvariant tract with variable functional effects and controversial clinical interpretation. It is located at the splice acceptor site at the end of intron 9 (in HGVS name; intron 8 in legacy name) and influences the amount of skipping of exon 10 (in the HGVS name; exon 9 in the legacy name). A CFTR mRNA without exon 10 originates a non-functional CFTR protein. The presence of a T5 repeat is usually considered necessary to originate pathological consequences. Although a small quote of anomalous splicing of the pre-mRNA CFTR may originate also from a (TG)mTn haplotype with a T repeat longer than the T5 (for example T7 and T9), the residual CFTR function assured in these cases by the high quote of correct splicing is considered sufficient to avoid clinical consequences. By contrast, the short T5 repeat (and even more so the very rare T3 repeat) is considered to produce a greater anomalous splicing and, consequently, to assure only a smaller quantity of correctly spliced CFTR mRNA [36]. As the Tn-dependent effect is modulated by the (TG)m repeat, the most frequent potentially pathological haplotypes are the (TG)11T5, (TG)12T5 and the (TG)13T5.

Our assay resulted in being able to measure from 1% to 99% of anomalously spliced CFTR mRNA (without exon 10) (consequently from 99% to 1% of wild-type CFTR mRNA). This dynamic range includes every value of the CFTR mRNA exon 10+/exon 10− ratio that is clinically significant [36,37]. In particular, it was revealed to be useful for the assessment of wild-type CFTR mRNA in the range from 25% to 12%, which is usually considered to include the threshold of residual functionality for the onset of mono-symptomatic CFTR-RD. Moreover, it was shown to be well suited to quantify the wild-type CFTR mRNA below 12%, up to 1%. In this range, the reduction of CFTR functionality corresponds to the onset of more severe forms of the disease. In particular, oligo-symptomatic CFTR-RD (consequent to a moderate reduction of functionality, near the upper end of the 12–1% range), poly-symptomatic forms of CF with pancreas sufficiency or poly-symptomatic forms of CF with pancreas insufficiency (consequent to a great reduction of functionality, near the lower end of the 12–1% range) may arise.

To check the proportion of correctly spliced CFTR mRNA produced by different (TG)mTn haplotypes, we selected homozygous (TG)mTn, eliminating the confounding effect of different (TG)mTn on the two alleles. The (TG)10T9 haplotype showed an average of 94.7% of correctly spliced CFTR mRNA, even in patients with both mutated alleles (Table 1, ID 1, 2 and 3). The (TG)10T7 haplotype originated in an average of 88.0% of correctly spliced CFTR mRNA as measured in a patient with both mutated alleles (Table 1, ID 4) and in a heterozygous subject (Table 1, ID 5). The (TG)11T7 haplotype originated an average of 80.9% of correctly spliced CFTR mRNA as measured in patients with both mutated alleles (Table 1, ID 6 and 7) and 3 heterozygous subjects (Table 1, ID 8, 9 and 10). Up to about 6% of variability in the amount of the spliced forms was shown by the (TG)10T9, (TG)10T7 and (TG)11T7 haplotypes. Our results confirm that, despite the variability, the amount of correctly spliced CFTR mRNA produced by these haplotypes without a T5 allele is always above the 77%, which is more than enough to prevent pathological consequences. By contrast, the (TG)11T5 haplotype analyzed in a subject with no other CFTR variants (Table 1, ID 11) was shown to have 47.5% of correctly spliced CFTR mRNA that, assuming an equal contribution of the two alleles (in the absence of other variants), would correspond to 23.8% of residual functionality from one allele. An overall CFTR functionality of 47.5% in a (TG)11T5/(TG)11T5 homozygous subject, with no other CFTR variants as in this case, is sufficient to prevent pathological manifestations. However, if on one allele there is a pathological CFTR variant with no residual CFTR function in compound heterozygosity with a (TG)11T5 allele, the overall residual functionality of 23.8% would be a value entering in the possible pathological area. This finding confirms our previous results [8] about the possibility that also the (TG)11T5 haplotype may cause some forms of CFTR-RD. The (TG)12T5 tract was found to produce low levels of correctly spliced CFTR mRNA, on average 27.6% with, however, a great variability between a heterozygous subject with 36.2% of correct splicing (Table 1, ID 12) and a subject with 19.1% of correctly spliced

CFTR mRNA and no other pathological variants (Table 1, ID 13). In particular for this last subject, assuming an equal allelic contribution in the absence of any other CFTR variants, the allelic residual functionality would be 9.6%. The high variability and low levels of residual functionality of the (TG)12T5 haplotype suggest the possibility, as also evidenced by our previous results [8], that the (TG)12T5 tract may sustain, in addition to CFTR-RD, also mild forms of CF.

In effect, it has already been demonstrated that the same (TG)mT5 haplotype may generate different amounts of correctly spliced CFTR mRNA in different cellular forms and tissues [38–40], as well as that the same (TG)mT5 haplotype can be linked to different clinical macrocategories [22]. In particular, the (TG)11T5 tract, that undergoes a moderate anomalous splicing and usually produces a moderate quantity of wild-type CFTR mRNA, has been found in CFTR-RD and CBAVD (Congenital Bilateral Absence of the Vas Deferens). The (TG)12T5 tract, that undergoes a high anomalous splicing associated to a low quantity of wild-type CFTR mRNA, has been found in CF with pancreas sufficiency, CFTR-RD and CBAVD. Finally, the (TG)13T5 tract, that undergoes a very high anomalous splicing linked to a very low amount of wild-type CFTR mRNA, has been found in CF with pancreas sufficiency and CFTR-RD.

The great variability in the functional and clinical effect of the (TG)mT5 haplotypes creates the need for a rapid method of quantification of the CFTR mRNA with and without exon 10, starting from clinical samples. In respect to other quantitative methods already proposed [41,42], our method has the advantage of being rapid, economical and easy to perform. In fact, it has the minimal technical requirements of a standard real time PCR assay, uses assays either disclosed or commercially available and allows a simplified calculation of the relative quantity of exon 10+/exon 10− ratio. The validation we performed on clinical samples by comparison with another method of evaluation allows direct use without further validation by other users, if our experimental conditions are reproduced.

The functional effect of (TG)mTn is variable, also depending on other factors additional to the haplotype. It is hard to predict the clinical manifestations only from the knowledge of the haplotype. Consequently, the experimental evaluation of residual wild-type CFTR mRNA is mandatory. Our method supports the enhancement of the evaluation of the functional effect of different (TG)mTn haplotypes, with an amelioration of the comprehension of the genotype–phenotype relationship in CF and CFTR-RD. The quantitative evaluation of residual wild-type CFTR mRNA, in particular for (TG)13T5, (TG)12T5 or (TG)11T5 haplotypes, is crucial for diagnosis, prognosis and personalized therapy. From a diagnostic point of view, it helps in defining whether the wild type CFTR mRNA amount is below the disease-causing threshold. Furthermore, the smaller the amount of exon 10+, the more severe the clinical manifestations will be, and also the prognostic ability will be improved. In particular, knowledge of the residual quantity of wild-type CFTR helps in defining if CFTR-RD, and mild or severe CF can be expected. Finally, in the era of precision medicine for CF, an accurate assessment of the amount of anomalously spliced CFTR mRNA may allow specific therapeutic interventions. Although there are still no specific treatments in clinical practice effective for CFTR splicing variants, aimed at restoring the correct splicing, several drugs are under study and/or undergoing clinical trials. When these treatments become available to CF patients, knowledge of the amount of CFTR spliced/unspliced forms from (TG)mT5 haplotypes will guide their correct application.

References

1. Lucarelli, M.; Porcaro, L.; Biffignandi, A.; Costantino, L.; Giannone, V.; Alberti, L.; Bruno, S.M.; Corbetta, C.; Torresani, E.; Colombo, C.; et al. A New Targeted CFTR Mutation Panel Based on Next-Generation Sequencing Technology. *J. Mol. Diagn.* **2017**, *19*, 788–800. [CrossRef] [PubMed]

2. Ferraguti, G.; Pierandrei, S.; Bruno, S.M.; Ceci, F.; Strom, R.; Lucarelli, M. A template for mutational data analysis of the CFTR gene. *Clin. Chem. Lab. Med.* **2011**, *49*, 1447–1451. [CrossRef] [PubMed]

3. Ivanov, M.; Matsvay, A.; Glazova, O.; Krasovskiy, S.; Usacheva, M.; Amelina, E.; Chernyak, A.; Ivanov, M.; Musienko, S.; Prodanov, T.; et al. Targeted sequencing reveals complex, phenotype-correlated genotypes in cystic fibrosis. *BMC Med Genom.* **2018**, *11*, 13. [CrossRef]

4. Bergougnoux, A.; D'Argenio, V.; Sollfrank, S.; Verneau, F.; Telese, A.; Postiglione, I.; Lackner, K.J.; Claustres, M.; Castaldo, G.; Rossmann, H.; et al. Multicenter validation study for the certification of a CFTR gene scanning method using next generation sequencing technology. *Clin. Chem. Lab. Med.* **2018**, *56*, 1046–1053. [CrossRef] [PubMed]

5. Bergougnoux, A.; Lopez, M.; Girodon, E. The Role of Extended CFTR Gene Sequencing in Newborn Screening for Cystic Fibrosis. *Int. J. Neonatal Screen.* **2020**, *6*, 23. [CrossRef] [PubMed]

6. Bergougnoux, A.; Taulan-Cadars, M.; Claustres, M.; Raynal, C. Current and future molecular approaches in the diagnosis of cystic fibrosis. *Expert Rev. Respir. Med.* **2018**, *12*, 415–426. [CrossRef]

7. Straniero, L.; Soldà, G.; Costantino, L.; Seia, M.; Melotti, P.; Colombo, C.; Asselta, R.; Duga, S. Whole-gene CFTR sequencing combined with digital RT-PCR improves genetic diagnosis of cystic fibrosis. *J. Hum. Genet.* **2016**, *61*, 977–984. [CrossRef]

8. Lucarelli, M.; Bruno, S.M.; Pierandrei, S.; Ferraguti, G.; Testino, G.; Truglio, G.; Strom, R.; Quattrucci, S. The Impact on Genetic Testing of Mutational Patterns of CFTR Gene in Different Clinical Macrocategories of Cystic Fibrosis. *J. Mol. Diagn.* **2016**, *18*, 554–565. [CrossRef]

9. Terlizzi, V.; Piccialli, G.; Salvatore, D.; Lucarelli, M.; Raia, V.; Angioni, A.; Carnovale, V.; Cirilli, N.; Casciaro, R.; Colombo, C.; et al. Genotype–phenotype correlation and functional studies in patients with cystic fibrosis bearing CFTR complex alleles. *J. Med. Genet.* **2017**, *54*, 224–235. [CrossRef]

10. Tsui, L.-C.; Dorfman, R. The Cystic Fibrosis Gene: A Molecular Genetic Perspective. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a009472. [CrossRef]

11. Cutting, G.R. Cystic fibrosis genetics: From molecular understanding to clinical application. *Nat. Rev. Genet.* **2015**, *16*, 45–56. [CrossRef] [PubMed]

12. Terlizzi, V.; Lucarelli, M.; Salvatore, D.; Angioni, A.; Bisogno, A.; Braggion, C.; Buzzetti, R.; Carnovale, V.; Casciaro, R.; Castaldo, G.; et al. Clinical expression of cystic fibrosis in a large cohort of Italian siblings. *BMC Pulm. Med.* **2018**, *18*, 196. [CrossRef] [PubMed]

13. Terlizzi, V.; Carnovale, V.; Piccialli, G.; Castellani, C.; Cirilli, N.; Colombo, C.; Corti, F.; Cresta, F.; D'Adda, A.; Lucarelli, M.; et al. Clinical expression of patients with the D1152H CFTR mutation. *J. Cyst. Fibros.* **2015**, *14*, 447–452. [CrossRef]

14. Narzi, L.; Ferraguti, G.; Stamato, A.; Narzi, F.; Valentini, S.; Lelli, A.; Delaroche, I.; Lucarelli, M.; Strom, R.; Quattrucci, S. Does cystic fibrosis neonatal screening detect atypical CF forms? Extended genetic characterization and 4-year clinical follow-up. *Clin. Genet.* **2007**, *72*, 39–46. [CrossRef] [PubMed]

15. Bareil, C.; Guittard, C.; Altieri, J.-P.; Templin, C.; Claustres, M.; Georges, M.D. Comprehensive and Rapid Genotyping of Mutations and Haplotypes in Congenital Bilateral Absence of the Vas Deferens and Other Cystic Fibrosis Transmembrane Conductance Regulator-Related Disorders. *J. Mol. Diagn.* **2007**, *9*, 582–588. [CrossRef] [PubMed]

16. Bobadilla, J.L.; Macek, M.; Fine, J.P.; Farrell, P.M. Cystic fibrosis: A worldwide analysis ofCFTR mutations?correlation with incidence data and application to screening. *Hum. Mutat.* **2002**, *19*, 575–606. [CrossRef] [PubMed]

17. Beauchamp, K.A.; Taber, K.A.J.; Grauman, P.V.; Spurka, L.; Lim-Harashima, J.; Svenson, A.; Goldberg, J.D.; Muzzey, D. Sequencing as a first-line methodology for cystic fibrosis carrier screening. *Genet. Med.* **2019**, *21*, 2569–2576. [CrossRef]

18. Bonadia, L.C.; Marson, F.A.L.; Ribeiro, J.D.; Paschoal, I.A.; Pereira, M.C.; Ribeiro, A.F.; Battagin, A.S. CFTR genotype and clinical outcomes of adult patients carried as cystic fibrosis disease. *Gene* **2014**, *540*, 183–190. [CrossRef]

19. Bear, C.E. A Therapy for Most with Cystic Fibrosis. *Cell* **2020**, *180*, 211. [CrossRef]

20. Amaral, M.D. Novel personalized therapies for cystic fibrosis: Treating the basic defect in all patients. *J. Intern. Med.* **2015**, *277*, 155–166. [CrossRef]

21. Rossi, T.; Grandoni, F.; Mazzilli, F.; Quattrucci, S.; Antonelli, M.; Strom, R.; Lucarelli, M. High frequency of (TG)mTn variant tracts in the cystic fibrosis transmembrane conductance regulator gene in men with high semen viscosity. *Fertil. Steril.* **2004**, *82*, 1316–1322. [CrossRef] [PubMed]

22. Lucarelli, M.; Bruno, S.M.; Pierandrei, S.; Ferraguti, G.; Stamato, A.; Narzi, F.; Amato, A.; Cimino, G.; Bertasi, S.; Quattrucci, S.; et al. A Genotypic-Oriented View of CFTR Genetics Highlights Specific Mutational Patterns Underlying Clinical Macrocategories of Cystic Fibrosis. *Mol. Med.* **2015**, *21*, 257–275. [CrossRef]

23. Claustres, M. Molecular pathology of the CFTR locus in male infertility. *Reprod. Biomed. Online* **2005**, *10*, 14–41. [CrossRef]

24. Stuhrmann, M.; Dörk, T. CFTR gene mutations and male infertility. *Andrologia* **2000**, *32*, 71–83. [CrossRef] [PubMed]

25. Pierandrei, S.; Blacona, G.; Fabrizzi, B.; Cimino, G.; Cirilli, N.; Caporelli, N.; Angeloni, A.; Cipolli, M.; Lucarelli, M. Two novel and correlated CF-causing insertions in the (TG)mTn tract of the CFTR gene. *PLoS ONE* **2019**, *14*, e0222838. [CrossRef]

26. Giordano, S.; Amato, F.; Elce, A.; Monti, M.; Iannone, C.; Pucci, P.; Seia, M.; Angioni, A.; Zarrilli, F.; Castaldo, G.; et al. Molecular and Functional Analysis of the Large 5′ Promoter Region of CFTR Gene Revealed Pathogenic Mutations in CF and CFTR-Related Disorders. *J. Mol. Diagn.* **2013**, *15*, 331–340. [CrossRef]

27. Tomaiuolo, A.C.; Alghisi, F.; Petrocchi, S.; Surace, C.; Roberti, M.C.; Bella, S.; Lucidi, V.; Angioni, A. Clinical hallmarks and genetic polymorphisms in the CFTR gene contribute to the disclosure of the A1006E mutation. *Clin. Investig. Med.* **2010**, *33*, E234–E239. [CrossRef]

28. Pagin, A.; Devos, A.; Figeac, M.; Truant, M.; Willoquaux, C.; Broly, F.; Lalau, G. Applicability and Efficiency of NGS in Routine Diagnosis: In-Depth Performance Analysis of a Complete Workflow for CFTR Mutation Analysis. *PLoS ONE* **2016**, *11*, e0149426. [CrossRef]

29. Kiesewetter, S.; Macek, M.; Davis, C.; Curristin, S.M.; Chu, C.-S.; Graham, C.; Shrimpton, A.E.; Cashman, S.M.; Tsui, L.-C.; Mickle, J.; et al. A mutation in CFTR produces different phenotypes depending on chromosomal background. *Nat. Genet.* **1993**, *5*, 274–278. [CrossRef]

30. Massie, R.; Poplawski, N.; Wilcken, B.; Goldblatt, J.; Byrnes, C.; Robertson, C. Intron-8 polythymidine sequence in Australasian individuals with CF mutations R117H and R117C. *Eur. Respir. J.* **2001**, *17*, 1195–1200. [CrossRef]

31. Feldmann, D.; Couderc, R.; Audrezet, M.-P.; Ferec, C.; Bienvenu, T.; Desgeorges, M.; Claustres, M.; Blayau, M.; Bozon, D.; Malinge, M.-C.; et al. CFTR genotypes in patients with normal or borderline sweat chloride levels. *Hum. Mutat.* **2003**, *22*, 340. [CrossRef] [PubMed]

32. Cuppens, H.; Lin, W.; Jaspers, M.; Costes, B.; Teng, H.; Vankeerberghen, A.; Jorissen, M.; Droogmans, G.; Reynaert, I.; Goossens, M.; et al. Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (Tg)m locus explains the partial penetrance of the T5 polymorphism as a disease mutation. *J. Clin. Investig.* **1998**, *101*, 487–496. [CrossRef] [PubMed]

33. Niksic, M.; Romano, M.; Buratti, E.; Pagani, F.; Baralle, F.E. Functional Analysis of cis-Acting Elements Regulating the Alternative Splicing of Human CFTR Exon 9. *Hum. Mol. Genet.* **1999**, *8*, 2339–2349. [CrossRef] [PubMed]

34. Lucarelli, M.; Grandoni, F.; Rossi, T.; Mazzilli, F.; Antonelli, M.; Strom, R. Simultaneous Cycle Sequencing Assessment of (TG)m and Tn Tract Length in CFTR Gene. *Biotechniques* **2002**, *32*, 540–547. [CrossRef]

35. Lucarelli, M.; Narzi, L.; Piergentili, R.; Ferraguti, G.; Grandoni, F.; Quattrucci, S.; Strom, R. A 96-well formatted method for exon and exon/intron boundary full sequencing of the CFTR gene. *Anal. Biochem.* **2006**, *353*, 226–235. [CrossRef]

36. Chillón, M.; Casals, T.; Mercier, B.; Bassas, L.; Lissens, W.; Silber, S.; Romey, M.-C.; Ruiz-Romero, J.; Verlingue, C.; Claustres, M.; et al. Mutations in the Cystic Fibrosis Gene in Patients with Congenital Absence of the Vas Deferens. *N. Engl. J. Med.* **1995**, *332*, 1475–1480. [CrossRef]

37. Zhang, L.; Button, B.; Gabriel, S.E.; Burkett, S.; Yan, Y.; Skiadopoulos, M.H.; Dang, Y.L.; Vogel, L.N.; McKay, T.; Mengos, A.; et al. CFTR Delivery to 25% of Surface Epithelial Cells Restores Normal Rates of Mucus Transport to Human Cystic Fibrosis Airway Epithelium. *PLoS Biol.* **2009**, *7*, e1000155. [CrossRef]

38. Buratti, E.; Brindisi, A.; Pagani, F.; Fe, B. Nuclear Factor TDP-43 Binds to the Polymorphic TG Repeats in CFTR Intron 8 and Causes Skipping of Exon 9: A Functional Link with Disease Penetrance. *Am. J. Hum. Genet.* **2004**, *74*, 1322–1325. [CrossRef]

39. Pagani, F.; Buratti, E.; Stuani, C.; Romano, M.; Zuccato, E.; Niksic, M.; Giglio, L.; Faraguna, D.; Baralle, F.E. Splicing Factors Induce Cystic Fibrosis Transmembrane Regulator Exon 9 Skipping through a Nonevolutionary Conserved Intronic Element. *J. Biol. Chem.* **2000**, *275*, 21041–21047. [CrossRef]

40. Pagani, F.; Baralle, F.E. Genomic variants in exons and introns: Identifying the splicing spoilers. *Nat. Rev. Genet.* **2004**, *5*, 389–396. [CrossRef]

41. Amaral, M.D.; Clarke, L.A.; Ramalho, A.; Beck, S.; Broackes-Carter, F.; Rowntree, R.; Mouchel, N.; Williams, S.H.; Harris, A.; Tzetis, M.; et al. Quantitative methods for the analysis of CFTR transcripts/splicing variants. *J. Cyst. Fibros.* **2004**, *3* (Suppl. 2), 17–23. [CrossRef] [PubMed]

42. Andrieux, J.; Audrézet, M.P.; Frachon, I.; Leroyer, C.; Roge, C.; Scotet, V.; Férec, C. Quantification of CFTR splice variants in adults with disseminated bronchiectasis, using the TaqMan fluorogenic detection system. *Clin. Genet.* **2002**, *62*, 60–67. [CrossRef] [PubMed]