Clinica Chimica Acta 558 (2024) 118317

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Identification of an ultra-rare Alu insertion in the *CFTR* gene: Pitfalls and challenges in genetic test interpretation

Speranza Esposito [a,b,1], Immacolata Zollo [a,b,1], Valeria Rachela Villella [a,b], Filippo Scialò [b,f], Sonia Giordano [c], Maria Valeria Esposito [d], Nunzia Salemme [e], Carmela Di Domenico [b], Gustavo Cernera [a,b], Federica Zarrilli [a,b], Giuseppe Castaldo [a,b], Felice Amato [a,b,*]

[a] Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II, Naples, Italy
[b] CEINGE- Advanced Biotechnologies Franco Salvatore, Naples, Italy
[c] AORN Ospedali dei Colli-Monaldi-Cotugno-CTO, Naples, Italy
[d] ASL Napoli 1 Centro, Naples, Italy
[e] San Giuseppe and Melorio Hospital, Santa Maria Capua Vetere, Caserta, Italy
[f] Department of Translational Medical Science, University of Campania Luigi Vanvitelli, Naples, Italy

ARTICLE INFO

Keywords:
CFTR
Cystic fibrosis
Genetic analysis
NGS
Appropriateness

ABSTRACT

Cystic fibrosis (CF) is a life-limiting genetic disorder characterized by defective chloride ion transport due to mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Early detection through newborn screening programs significantly improves outcomes for individuals with CF by enabling timely intervention. Here, we report the identification of an Alu element insertion within the exon 15 of CFTR gene, initially overlooked in standard next-generation sequencing analyses. However, using traditional molecular techniques, based on polymerase chain reaction and Sanger sequencing, allowed the identification of the Alu element and the reporting of a correct diagnosis. Our analysis, based on bioinformatics tools and molecular techniques, revealed that the Alu element insertion severely affects the gene expression, splicing patterns, and structure of CFTR protein. In conclusion, this study emphasizes the importance of how the integration of human expertise and modern technologies represents a pivotal step forward in genomic medicine, ensuring the delivery of precision healthcare to individuals affected by genetic diseases.

1. Introduction

Cystic fibrosis (CF) is an autosomal recessive disease caused by mutations in the CF transmembrane conductance regulator gene (CFTR). Since its identification in 1989 [1], more than 2100 variants have been identified and recorded in the CF Mutation Database (https://www.genet.sickkids.on.ca/StatisticsPage.html). This has been possible thanks to the improvement of DNA sequencing technologies that since the development of Sanger sequencing in 1977 [2], in only half century have evolved into the most sophisticated Next-generation sequencing (NGS) platforms [3].

Here we report the identification of an Alu insertion in the CFTR gene, within exon 15. To date, this would be the third case of an Alu sequence insertion in the CFTR gene. The first two Alu insertions were reported in 2008 [4]. In this work, J.M. Chen and colleagues screened 100 previously unresolved CF chromosomes using quantitative high-performance liquid chromatography, identifying two Alu insertions in exons 16 and 17b of the CFTR gene.

Advances in high-throughput sequencing technologies and bioinformatics tools have enabled the identification and characterization of Alu insertions on a genome-wide scale.

Alu elements are DNA sequences approximately 300 base pairs in length and are characterized by a conserved internal promoter region that drives transcription of their RNA intermediates [5]. Among the various types of Mobile Genetic Elements (MGEs), Alu elements are the most abundant in the human genome, comprising approximately 11 % of the total DNA. The MGEs are essential components of genomic diversity and have played a crucial role in shaping the evolution of genomes. The transposition of Alu elements is primarily mediated through the enzymatic activity of the Long Interspersed Elements (LINE)-1 (L1)

* Corresponding author at: University of Naples Federico II, Italy.
E-mail address: felice.amato@unina.it (F. Amato).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cca.2024.118317
Received 14 March 2024; Received in revised form 28 March 2024; Accepted 3 April 2024
Available online 4 April 2024

retrotransposon machinery leading to their proliferation throughout the human genome [6].

Alu insertions can result in various genomic rearrangements, including duplications, deletions, inversions, and chromosomal fusions. These structural variations can alter gene expression patterns and contribute to genomic diversity between individuals and populations. Alu insertions have been implicated in the creation of novel genes and exons, alternative splicing events, and the generation of genetic diversity that may be subject to natural selection [7].

Furthermore, Alu insertions have been linked to several human diseases. For instance, Alu insertions have been implicated in neurodegenerative disorders, including Alzheimer's disease [8,9] and Parkinson's disease [10], and in genetic disorders such as haemophilia [11] and neurofibromatosis [12].

In conclusion, although next-generation sequencing has revolutionized genetic analysis and has significantly improved the diagnosis of CF with a sensitivity close to 99 %, the understanding of novel and ethnic-specific variants is still challenging, especially for newborn and CF carrier screening, or for all CF related conditions [13–15]. Furthermore, with this work, we would like to emphasize the role of the human factor in interpreting and reporting the genetic analysis data, highlighting the crucial role of highly specialized and expert personnel versus the bioinformatic output of NGS data. For this purpose, we present the identification of this Alu insertion into exon 15 of CFTR gene, as an example that illustrates a pitfalls that prevented NGS alone from providing a correct diagnosis.

2. Results and discussion

2.1. CFTR gene next-generation sequencing analysis

A 5T/12TG allele was identified in the CFTR gene as a result of the genetic analysis of a newborn with altered trypsinogen values. As shown in Fig. 1A, we identified also a likely deletion in the exon 15. The gold standard MLPA assay was carried out to confirm the likely deletion [16]. This assay foresees the use of probes binding specific genomic regions that, after a ligation process, are amplified by PCR, separated by capillary electrophoresis, and analysed (Fig. 1B and C). Surprisingly, the MPLA assay did not confirm the NGS results because no deletion was identified (Fig. 1C). The case would have been reported as positive only for the presence of the 5T/12TG allele.

However, the same exon15 likely deletion was also detected in one of the analysed parents. At this point, given the relationship between the two analysed samples, we hypothesized that the warning shown by the NGS analysis could be caused by some kind of DNA sequence alteration. Therefore, we performed a PCR amplification of the region of interest to better characterize the likely deletion. We got two PCR amplification products using the genomic DNA of one of our samples as template (Fig. 2). Contrary to what we expected, the smaller band corresponded to the wild-type sequence, whereas the amplification from the mutated allele generated a larger band, indicating the presence of an insertion. The blasting of the Sanger sequence of the upper PCR band with the NCBI nucleotide database revealed the insertion of the AluYb9 element. In particular, as seen in Fig. 2, the insertion consists of a total of 171-bp sequence corresponding to a 108-bp AluYb9 element plus an upstream 53-bp poly-T tail flanked by 10-bp direct repeats on both sides.

In conclusion, what appeared to be a deletion from NGS analysis is actually a 171 bp insertion. The reason why the MLPA analysis failed to identify any alteration in that region of interest is shown in the lower part of Fig. 2. The recognition sites of the Right (RPO) and Left (LPO) MLPA probes, highlighted in red and yellow respectively, are located just upstream of the Alu element insertion site, which in this way does not affect the hybridization of the MLPA probes. The deletion highlighted by the NGS analysis, instead, is due to the fact that the Alu

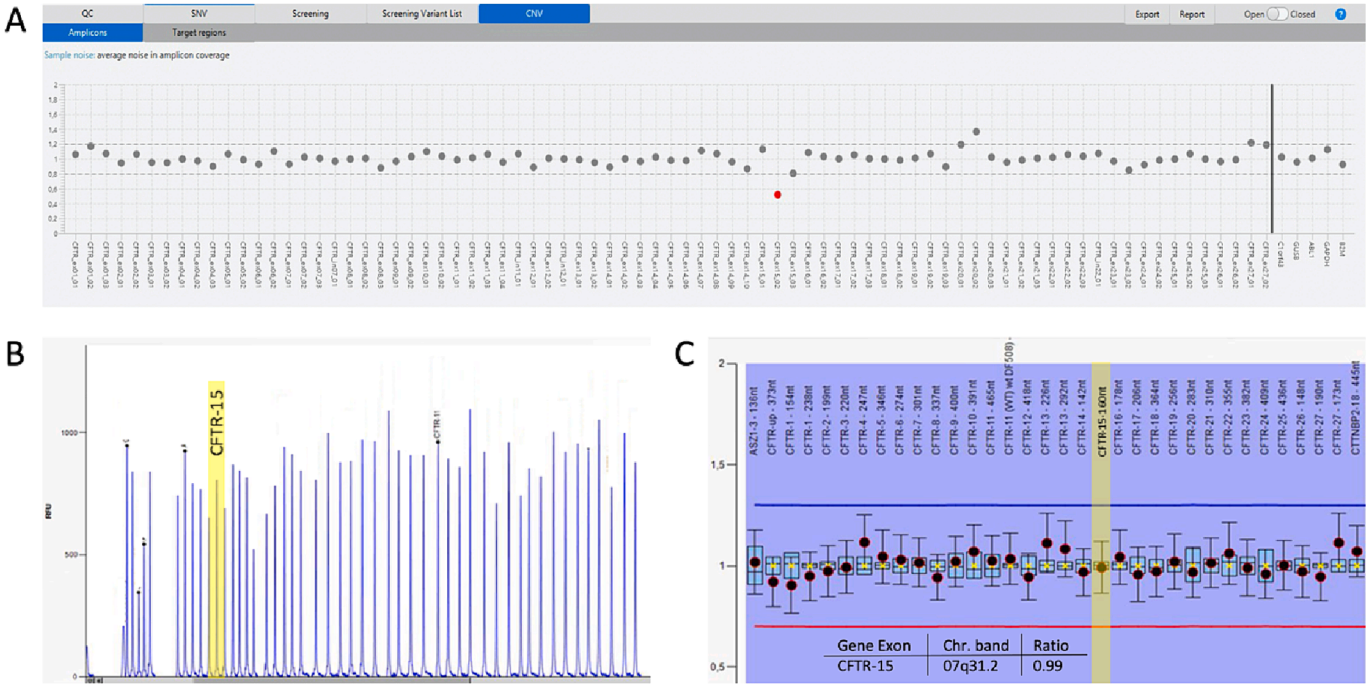

Fig. 1. Diagnostic Analysis. **A)** NGS output. The NGS analysis employed the Devyser CFTR kit which is based on the multiplex PCR amplification to create a library that covers the full target region in an overlapping way. The kit allows the detection of single nucleotide variations (SNVs), indels and copy number variations (CNVs) in the coding regions and adjacent exon–intron boundaries. In addition, it detects deep intronic mutation in introns 7, 11, 12, and 22. Each *CFTR* exon is amplified in different amplicons (grey dots) and their number depends on the exon size. In particular, the NGS output here reported shows for CNV analysis a warning at the level of amplicon 2 in the exon 15 (red dot), suggesting a likely deletion event. **B)** MLPA results. This analysis is used for the detection of duplications or deletions in a gene. PCR products are separated by capillary electrophoresis. Each fragment corresponds to a specific MLPA probe. **C)** The Final Ratio (FR), obtained by comparing the relative peak heights of reference probes and target probes in the test samples with those in reference samples with a known normal copy number, should be between 0.8 and 1,2. In the case of exon15 (highlighted in yellow) the FR is 0,99, thus indicating that no deletion is present in this region as in the entire gene.

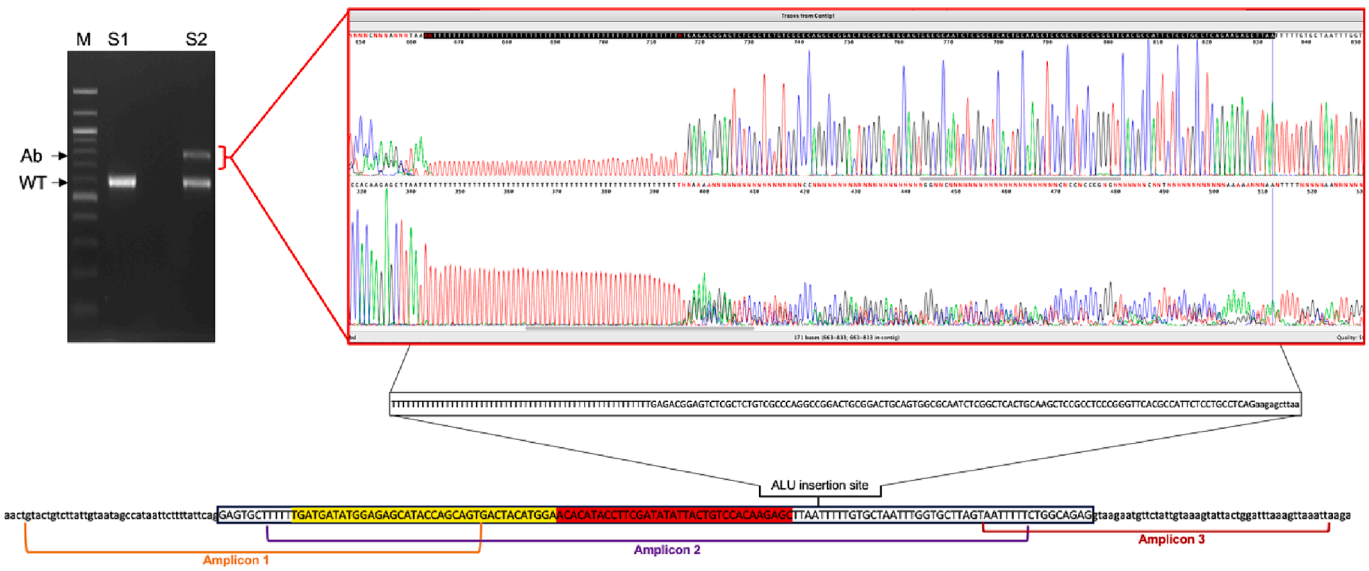

Fig. 2. PCR amplification and Sanger sequence analysis. The amplification of *CFTR* exon 15 results in one band (WT) from genomic DNA of a control subject (S1) and two bands from the patient subject (S2), the WT expected band, corresponding to the wildtype sequence, and an aberrant band (Ab). The Sanger sequence analysis of Ab band shows an insertion of a 171 bp long sequence. The exon 15 sequence is boxed and reported in uppercase letters with its intronic flanking regions in lowercase. The three amplicons obtained by NGS analysis are indicated. Also, the Right (red) and Left (yellow) Probes used in MLPA analysis (MLPA-RPO and MLPA-LPO respectively) are indicated. The insertion, downstream of the target region of MLPA probes, determined the failure of amplicon 2 amplification.

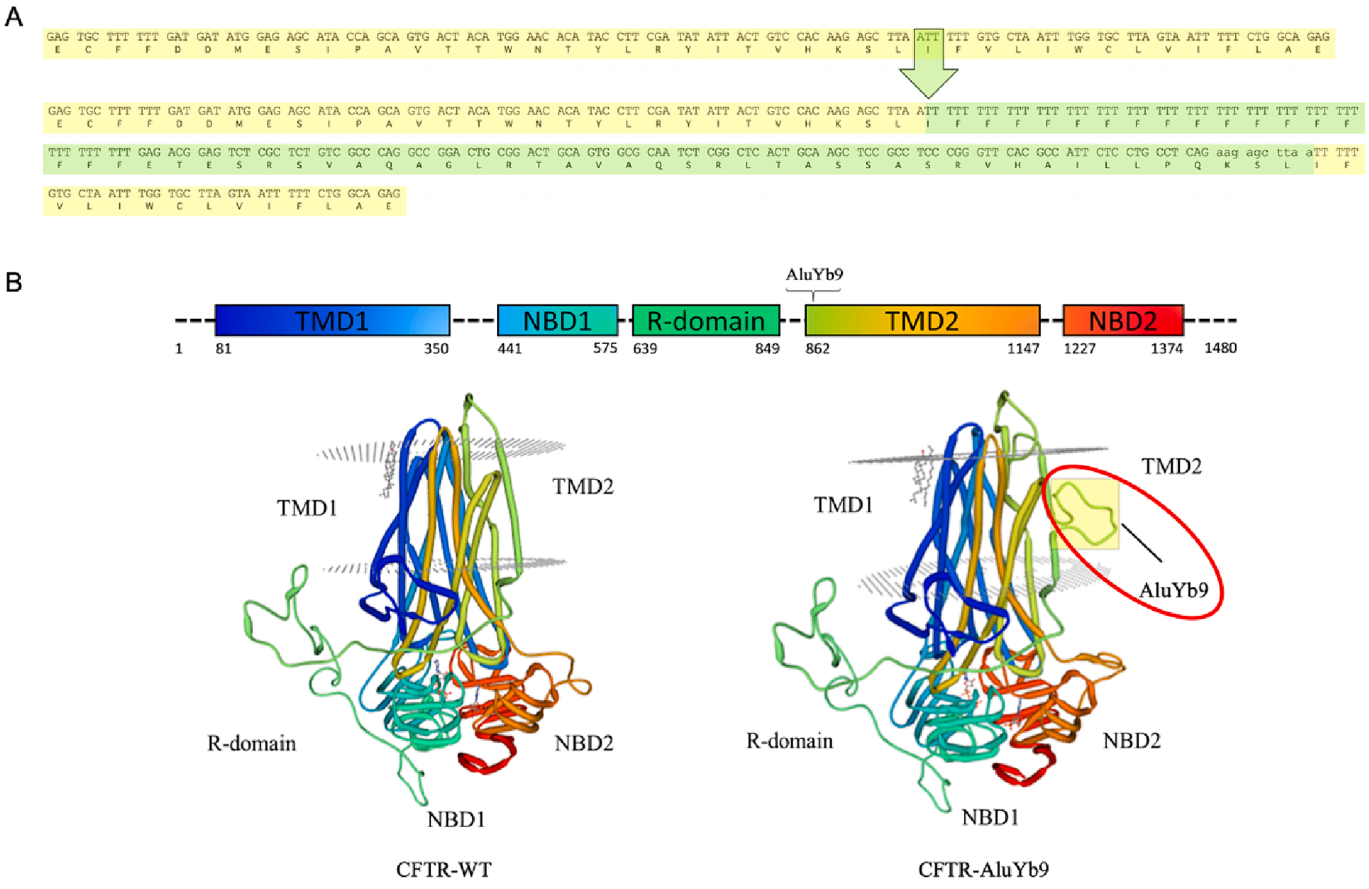

Fig. 3. *Swiss-model in silico prediction.* A) Nucleotide and protein sequence of CFTR exon 15 (light yellow) and Alu element (light green), and the Alu insertion site indicated by the arrow. B) Schematic representation of CFTR protein and its main protein domains (TMD for TransMembrane Domain 1 and 2, NBD for Nucleotide Binding Domain 1 and 2, and R for the regulatory domain. Structural model of CFTR WT (on the left) and CFTR-Alu (on the right) by Swiss-model. According to the *in silico* prediction, the insertion of the Alu element in the first alpha-helix of TMD2 causes a complete loss of folding of this domain.

element falls inside the amplicon 2, probably resulting in a poor or no amplification of this amplicon, and therefore in an incorrect deletion signal due to a lower copy number variation in that region.

3. Functional characterization of CFTR-Alu insertion

3.1. *In silico analyses*

To fully comprehend the impact that Alu insertion has on CFTR expression, processing, folding and membrane translocation, we first ran an *in silico* analysis using both Swiss-model software for protein structure analysis and NetGene2 for splice pattern analysis.

The Swiss-model software allowed us to reconstruct a 3D structure of CFTR-Alu and understand if and how this insertion could affect CFTR structure (Fig. 3B). Comparing the two models it can be seen that AluYb9 is located in the first alpha-helix of the transmembrane domain 2 (TMD2), corresponding precisely between the amino acids I860 and F861. In particular, the 171 bp of in-frame nucleotide insertion is (should be) translated in a 57 aa with a poly-phenylalanine stretch 17 aa long (Fig. 3A). It is reasonable to think that this type of insertion will destroy the helix-helix interactions with the consequent alteration of the tertiary structure of the transmembrane domains [17]. Most likely, during the insertion of the protein into the membrane, when individual transmembrane segments are placed into the membrane by the translocon machinery, the whole protein, due to its misfolded domain, could be recognized by the endoplasmic reticulum-associated protein degradation (ERAD) system to be destroyed by the cytoplasmic ubiquitin–proteasome system (UPS) [18–20].

On the other hand, to evaluate the effect of Alu insertion at the transcriptional level, we used the online available NetGene2 service for splicing pattern analysis [21,22]. Fig. 4 summarizes the results of this prediction, showing how the presence of the Alu element completely abolishes the canonical 3′splice acceptor site of exon 15 resulting probably in the skipping of the entire exon 15.

3.2. *Functional evaluation of recombinant CFTR protein*

To confirm both *in silico* predictions, we first inserted the Alu sequence into a lentiviral construct containing the CFTR and the yellow fluorescent protein (YFP) coding sequences separated by a T2A self-cleaving peptide. The obtained lentiviral particles bearing the expression cassette for Wild-type CFTR, F508del-CFTR, and Alu-CFTR were used to generate HEK293 cells stably expressing the corresponding CFTR proteins. The channel functionality of the CFTR proteins was assessed by the halide-sensitive YFP (HS-YFP) assay. As expected, the Alu-CFTR protein did not show any channel activity after forskolin stimulation in both treated and not treated cells with the triple Tezacaftor/Elexacaftor/Ivacaftor (TEI) combination (Fig. 5A and B).

In addition, as shown in (Fig. 5D), the western blot analysis demonstrates that the Alu-CFTR protein is not even produced. Since all the constructs with the various forms of CFTR support a high expression at the mRNA level (Fig. 5C), the lack of presence of the Alu-CFTR protein could be due, as predicted by the *in silico* analysis, to its misfolded structure that is recognized by the ERAD system and degraded by the cytoplasmic ubiquitin–proteasome system.

3.3. *Splicing pattern analysis using the minigene system*

As the expression of a protein via a plasmid construct does not take into account the possible effects that any genetic variation may have at the transcriptional level, we have cloned the exon15 WT or exon15-Alu with part of the upstream and downstream intronic regions, thereby conserving as much of the genomic context as possible, in the pMGene construct (Fig. 6A–B).

The experimental data confirmed the *in silico* prediction regarding the loss of the CFTR exon 15 acceptor site caused by the presence of the Alu sequence, causing the skipping of the entire exon 15. Instead, the predicted acceptor sites within the Alu sequence (Fig. 4), did not produce any alternative splicing product which includes part of exon 15.

Of course, as previously demonstrated also by our research group, analyzing messenger RNA directly from patient-derived cells is a gold standard method for studying splicing alterations in a correct cellular context, nasal, bronchial, or intestinal epithelium in the case of cystic fibrosis [23–26]. Since it considers the cell/tissue specificity of the splicing pattern and above all, the fact that regulatory elements necessary for the correct splicing can be localized in genomic regions that may be missing in the DNA region chosen to be cloned into the minigene plasmid. However, it should be underlined that the procedures for collecting these kinds of samples are relatively invasive for the patient. Even considering the nasal epithelium cells, with the less invasive

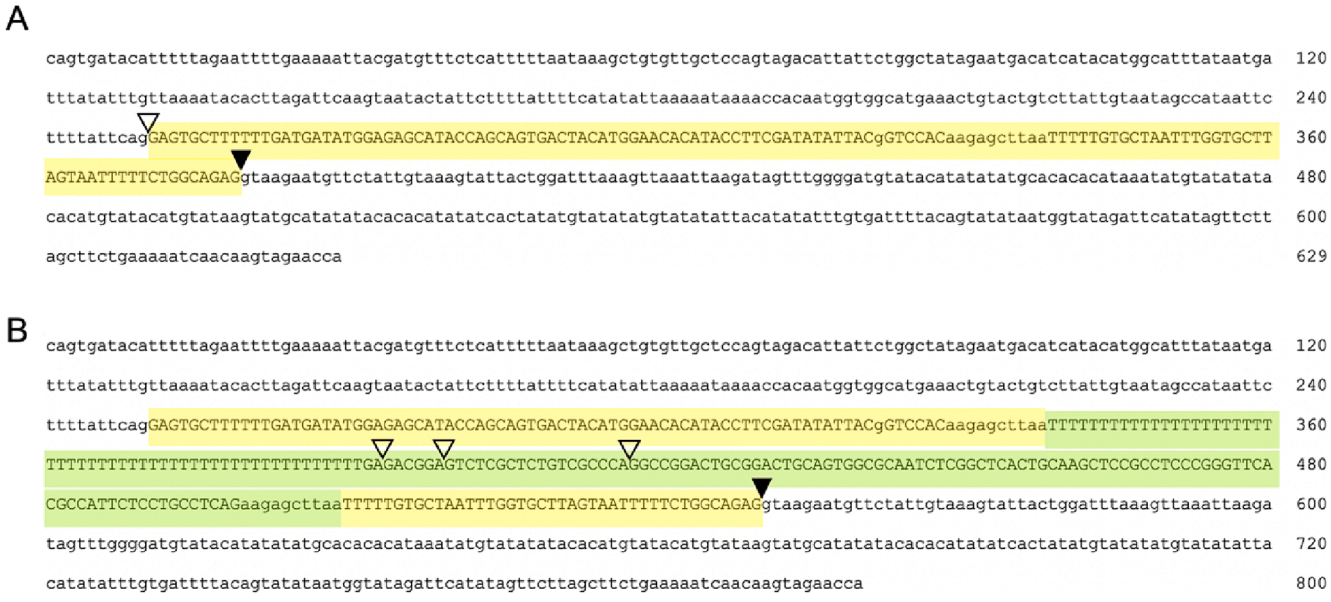

Fig. 4. NetGene2 *in silico* prediction. **A)** Nucleotide sequence of CFTR exon 15 (light yellow) with 250 bp intronic sequence up and downstream, **B)** same sequence in A) plus the Alu element (light green) included. The empty and filled triangles indicate the predicted acceptor and donor splice sites, respectively. In B, the software recognized three putative acceptor splice sites, all in the Alu element.

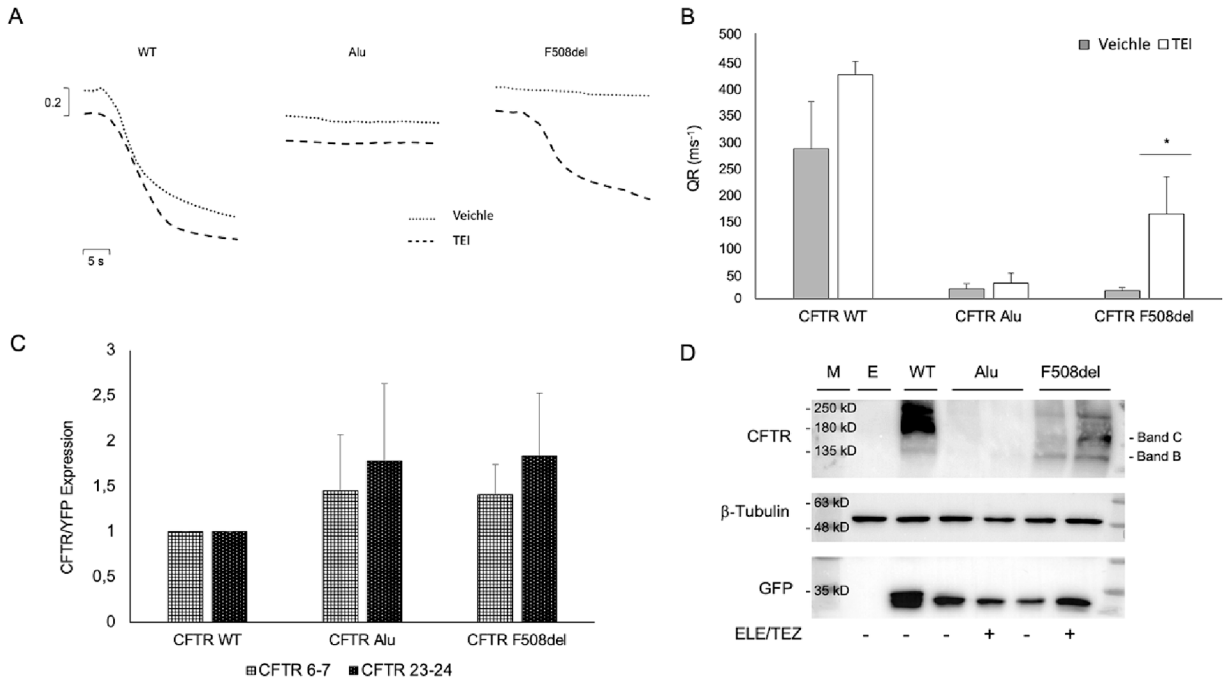

Fig. 5. Functional evaluation of recombinant CFTR protein. A and B) Representative traces and summary of data of the halide-sensitive yellow fluorescent protein (HS-YFP) assay on HEK293 cells, stably expressing the various forms of CFTR protein, treated or not with Elexacaftor (3 μM), Tezacaftor (3 μM), VX-770 (100 nM). Stimulation assay conditions: Forskolin (FSK, 20 μM). Each error bar reports the quenching rate (QR, mean, and StDv) calculated from three separate experiments in quadruplicate. *$p < 0.01$. C) CFTR mRNA expression analysis by qPCR assay using two different primer pairs, spanning between the CFTR exon 6–7 and 23–24 respectively. The CFTR expression was normalized for YFP co-expression. Each error bar reports mean and StDv calculated from three separate experiments D) Evaluation of CFTR protein expression and maturation by immunoblot analysis on lysates of HEK293 cells used for HS-YFP assay. M: protein marker. E: empty vector. WT, Alu and F508del-CFTR protein.

sampling procedure, it may still be poorly tolerated by very young patients or simply not permitted by parents. Thus, even considering these minigene assay limitations, it remains a valid alternative procedure to the patient's biological sample for studying the splicing pattern.

4. Materials and methods

4.1. Cell culture

Human embryonic kidney cells (HEK293 and HEK293T) were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % (v/v) fetal bovine serum (FBS), 100 U mL − 1 of penicillin–streptomycin (Pen-Strep), 1 mM of sodium pyruvate and 2 mM of L-glutamine (all from Euroclone, Milan, Italy) in a 37 °C, 5 % CO2 humidified incubator.

4.2. Molecular cloning

4.2.1. Generation of the minigene construct

To study the effect of the Alu insertion on the splicing pattern, the exon15 WT or exon15-Alu with part of the upstream and downstream intronic regions, was cloned in pMGene vector by homologous recombination using the M13 sequences. This vector contains the entire β-globin gene and a KpnI restriction site to clone the region of interest in the middle of intron 2 of β-Globin [27,28]. In particular, the region of interest was amplified from the heterozygous subject genomic DNA by PCR using the Velocity DNA Polymerase (#BIO-21098, Bioline) and the following primers: For: GTAAAACGACGGCCAGAAACACAAATG AGCTTTCAGTCT Rev: CAGGAAACAGCTATGACCGTGGTTCTACTTGT TG. The PCR program was set up as follows: initial denaturation stage 98 °C 2 min, 35 amplification cycles with three steps − 98 °C 5 s (denaturation), 56 °C 30 s (annealing), 72 °C 30 s (extension) - and a final extension at 72 °C for 7 min. The two bands (wild-type and Alu containing sequence) obtained by PCR were purified from the agarose gel using the NucleoSpin™ Gel and PCR Clean-up Kit (#740609, Macherey-Nagel™) according to the manufacturer's instructions. Then, they were inserted into pMGene vector by homologous recombination using NEBuilder HiFi DNA Assembly Cloning kit (#E5520, New England Biolabs). After Sanger sequencing to control the correct insertion of the region of interest, the two recombinant plasmids were later transfected in HEK293 cells by Attractene Transfection Reagent (#301005, QIAGEN). After 48 h, cells were collected and total RNA was extracted to be retrotranscribed for PCR analysis and Sanger analysis.

4.2.2. Generation of the Alu-CFTR expression construct

To generate the construct for the expression of CFTR bearing the Alu sequence into exon 15, the region of interest was amplified from the patient genomic DNA using the Velocity DNA Polymerase (#BIO-21098, Bioline) and the primers 3 and 4 of Table 1. Moreover, a 452 bp long fragment of exon 14, upstream the exon 15, and a 1019 bp long fragment downstream to exon 15 spanning from exon 16 to 22, were amplified using our construct with wild-type CFTR coding sequence as template, and the primers 1,2 and 5,6 respectively. All the primers were designed to produce overlapping PCR fragments to be cloned by homologous recombination using NEBuilder HiFi DNA Assembly Cloning kit (#E5520, New England Biolabs) in the lentiviral vector previously digested by EcoRI to excise the fragment to be replaced (Fig. 7).

4.3. Lentiviral packaging and generation of HEK293 cells expressing the WT-CFTR, F508del-CFTR or Alu-CFTR

The recombinant plasmids (YFP-CFTR-WT, YFP-CFTR-F508del, and YFP-CFTR Alu) were used to transfect HEK293T cells for the production of lentiviral particles. Briefly, $5x10^6$ cells were seeded in a 10 cm dish and transfected with a mix containing 2,5μg of pRSV-REV (#12253, Addgene), 3,5μg of pMD2.G (#12259, Addgene), 6,5 μg of pMDLg/

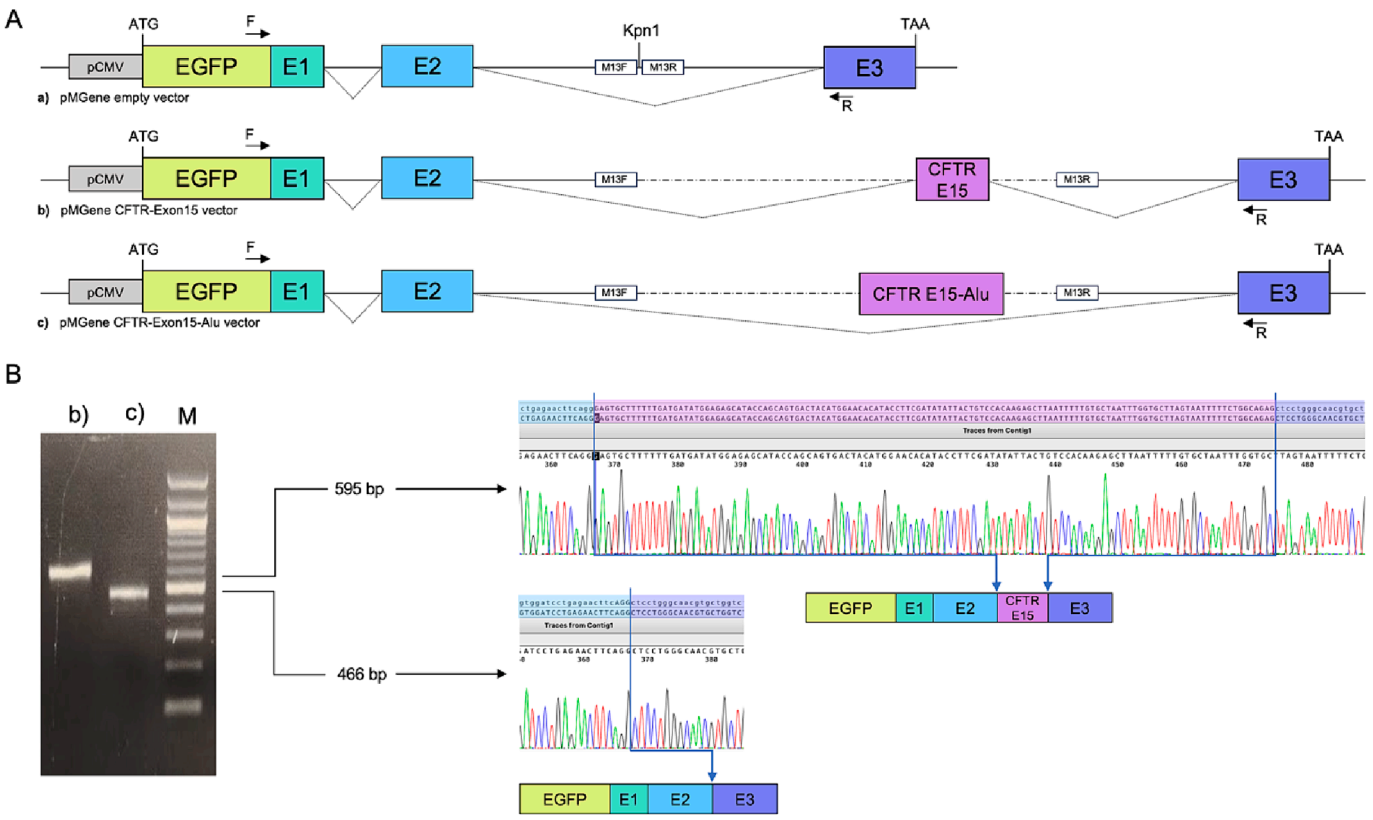

Fig. 6. Minigene assay A) Schematic structure of the pMGene constructs. EGFP is for Enhanced Green Fluorescent Protein, E1, E2, and E3 represent the β-globin gene exons (E1, E2, and E3), and CFTR 15 exon is also shown. Continuous lines between exons correspond to β-globin introns; whereas the dotted lines to CFTR introns. The cytomegalovirus promoter (pCMV) is shown in a thin box. Finally, the arrows F and R represent the primers used for the splicing pattern analysis. **B**) RT-PCR analysis of mRNA extracted from cell lines transfected with the indicated pMGene constructs. The Sanger sequences of the PCR products are in the right part of panel B, showing the complete skipping of CFTR exon 15 caused by the insertion of the Alu sequence.

Table 1
List of primers used to insert the Alu sequence into CFTR coding sequence.

1	CFTR-Ex14-For	5′- CAGACTGGAGAGTTTGGGGAAAAAAGGAAG −3′
2	CFTR-Ex15-5′-Rev	5′- CTCTCCATATCATCAAAAAAGCACTCC −3′
3	CFTR-Ex15-For	5′- GGAGTGCTTTTTTGATGATATGGAGAG −3′
4	CFTR-Ex15-Rev	5′- ACCTCTGCCAGAAAAATTACTAAGCACC −3′
5	CFTR-Ex15-3′-For	5′- GGTGCTTAGTAATTTTTCTGGCAGAGGT −3′
6	CFTR-Ex22-Rev	5′- CCAGATGTCATCTTTCTTCACGTGTGAAT −3′

pRRE (#12251, Addgene) and 10 μg of the transfer plasmid. The plasmid solution was made up to a final volume of 450 μl with 0,1XTE and 50 μl of 2,5M CaCl2 were added. The medium was replaced after 6 h and lentiviral particles collected 48 h after transfection. Later, lentiviral particles were used to transduce HEK293 cells.

4.4. RNA extraction and qRT-PCR

Total RNA was extracted using Direct-zol™ RNA MicroPrep Kit (#R2060, Zymo Research) according to the manufacturer's instructions. One μg of total RNA from each sample was used to obtain double-strand cDNA with SensiFAST™ cDNA Synthesis Kit (#BIO-65053, Bioline) setting up the following reaction: primer annealing at 25 °C for 10 min, reverse transcription at 42 °C for 15 min, inactivation at 85 °C for 5 min and final hold at 4 °C. Quantitative Real-Time PCR (qRT-PCR) was performed with the The StepOnePlus™ Real Time System instrument (Applied Biosystem).

For each PCR reaction, 10 μl 2x SensiFAST™ SYBR® Hi-ROX Kit (#BIO-92005, Bioline), 0,8μM of each primer (listed in Table 2) and 20 ng of the cDNA were used.

The reaction was set up as follows: initial holding stage 95 °C 2 min, 40 amplification cycles with three steps – 95 °C 5 s (denaturation), 60 °C 30 s (annealing), 95 °C 15 s, 60 °C 15 s, 95 °C 15 s (dissociation stage).

4.5. Western blot

Cells were lysed in NP40 lysis buffer (#FNN0021, Invitrogen) supplemented with protease inhibitors (#P8340, Sigma) and phosphatase inhibitors (#PIC2850). Soluble protein extracts were separated by SDS-polyacrylamide gel electrophoresis and transferred to PVDF membranes 0,2μm pore size (Amersham™) by elettroblotting. Membranes were blocked with 5 % non-fat dry milk and incubated with primary antibodies (anti-CFTR596: 1:1000, CFTR Antibody Distribution Program, Cystic Fibrosis Foundation, UNC-Chapel Hill, β-Tubulin-HRP: 1:1000, Abcam, ab21058, anti-GFP: 1:1000, Cell Signaling, #2956) ON Membranes were later incubated with horseradish peroxidase (HRP)-conjugated secondary antibody (1:8000 in 5 % non-fat dry milk) for 1 h at RT. Signals were finally detected with a chemiluminescent detection system (#E-IR-R301, Elabscience) by using Chemidoc MP Imaging System Biorad.

4.6. HS-YFP assay

HEK293 cells stably expressing CFTR-WT or CFTR-Alu together with a halide-sensitive yellow fluorescent protein (HS-YFP) were seeded onto clear-bottom 96-well black microplates. Where indicated, cells were treated with CFTR modulators - VX-661 (3 μM) + VX-445 (3 μM) + VX-770 (100 nM) - or vehicle (DMSO) for 24 h. Cells were rinsed two times with 200 μl of PBS (137 mM NaCl, 2.7 mM KCl, 8.1 mM Na2HPO4, 1.5

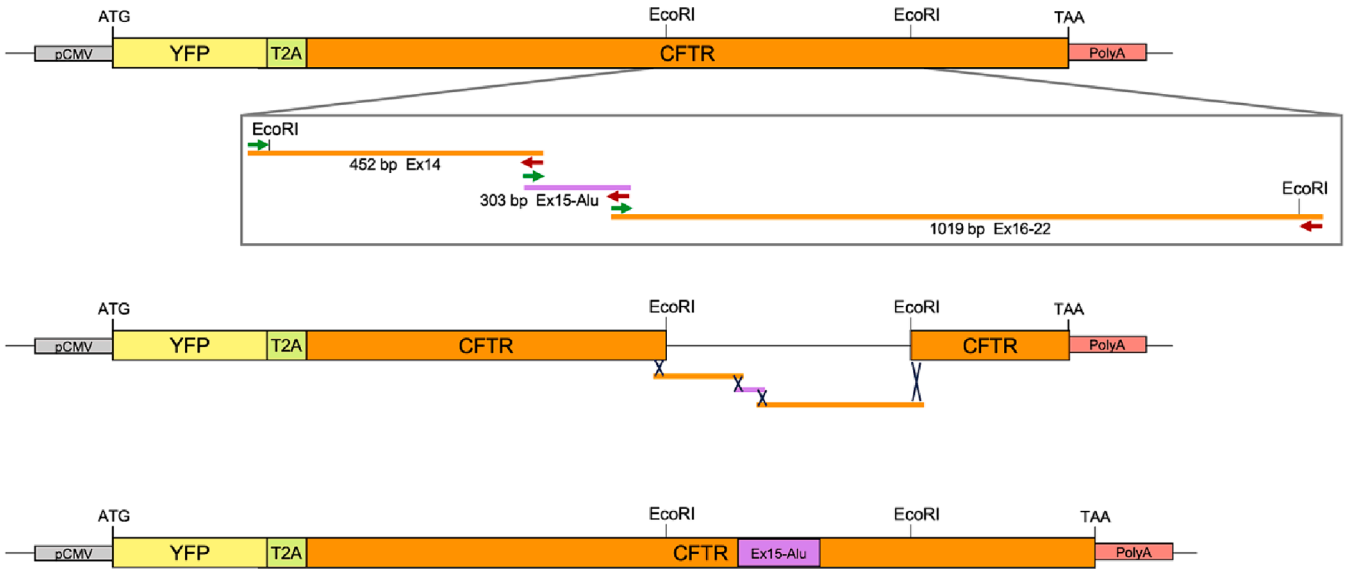

Fig. 7. Schematic representation of the strategy used to insert the Alu sequence into CFTR coding sequence.

Table 2
List of primers used for the study.

CFTR-Ex6-7 Fw	5′ - GTGGACTTGGTTTCCTGATAGT – 3′
CFTR-Ex6-7 Rev	5′ - ACTGATCTTCCCAGCTCTCT – 3′
CFTR-Ex23-24 Fw	5′ - GAGGAAAGCCTTTGGAGTGATA – 3′
CFTR-Ex23-24 Rev	5′ - CCACTGTTCATAGGGATCCAAG – 3′
YFP Fw	5′ - CAACAGCCAGAACGTCTATCTGA – 3′
YFP Rev	5′ - ATGTTGTGGCGGATCTTGAAG – 3′
GAPDH Fw	5′ - GAAGGTGAAGGTCGGAGTC – 3′
GAPDH Rev	5′ - GAAGATGGTGATGGGATTTC – 3′

mM KH2PO4, 1 mM CaCl2, and 0.5 mM MgCl2) and incubated for 30 min with 60 μl of the same solution with the addition of forskolin (20 μM). After incubation, 96-well microplates were transferred to a PHERAStar FSX plate reader. For each well, cell fluorescence was continuously measured for 5 s before and 20 s after injection of 165 μl of a modified PBS containing 137 mM NaI instead of NaCl (final I- concentration in the well: 100 mM). After background subtraction, cell fluorescence recordings were normalized for the initial value measured before addition of I- in each well. The signal decay caused by HS-YFP fluorescence quenching was fitted with a double exponential function to derive the maximal quenching rate (QR) that corresponds to initial influx of I- into the cells.

4.7. Functional splicing assays

The genomic region of interest, the exon 15 wt or exon15-Alu with part of the upstream and downstream intronic regions, was cloned in pMGene vector by homologous recombination. This vector contains the entire β-globin gene and a KpnI restriction site to clone the region of interest in the middle of intron 2 of β-Globin. The two recombinant plasmids were later transfected in HEK293 cells by Attractene Transfection Reagent (#301005, QIAGEN). After 48 h, cells were collected and total RNA extracted. After the *retro*-transcription, the cDNA obtained was used as template to set up a PCR to amplify the mRNA of interest.

4.8. In silico analyses

The *in silico* prediction of the effect of Alu element on splicing pattern was performed using the online available software https://services.healthtech.dtu.dk/services/NetGene2-2.42/., and the Swiss-model software for protein structure evaluations https://swissmodel.expasy.org/.

5. Conclusion

This is the third case reported in the literature of an insertion of the Alu element in the *CFTR* gene. As previously reported for the other two cases [4], also the Alu insertion here described causes a severe phenotype on the expression of the CFTR protein, confirmed by both *in silico* and experimental validations. However, this severe effect risked not being properly identified and reported, if the diagnosis had been based only on the current analytical technologies, such as next-generation sequencing. There is no doubt that the advent of NGS analysis and the increasingly use of artificial intelligence (AI) have significantly improved genetic analyses. However, despite the high accuracy and consistency of AI in tasks where the rules are well-defined and the data is clean, humans excel in tasks that require creative thinking, intuition, and adaptation to new situations. In summary, while AI can outperform humans in certain specific tasks in terms of speed, accuracy, and cost-effectiveness, humans still possess unique cognitive abilities such as creativity and adaptability that make them indispensable in many tasks. The ideal scenario often consists in exploiting the strengths of both humans and artificial intelligence to achieve optimal efficiency and effectiveness, particularly when decisions must be made about the well-being of patients, both from a therapeutic and diagnostic point of view as in this case.

Funding

PRIN Progetti di Ricerca di Rilevante Interesse Nazionale 2022. “Theratyping of Cystic Fibrosis”. Codice progetto n. 2022FRSS2H/ Ministero dell’Università e della Ricerca.

CRediT authorship contribution statement

Speranza Esposito: Investigation. **Immacolata Zollo:** Investigation. **Valeria Rachela Villella:** Investigation. **Filippo Scialò:** Methodology, Investigation. **Sonia Giordano:** Investigation. **Maria Valeria Esposito:** Investigation. **Nunzia Salemme:** Investigation. **Carmela Di Domenico:** Investigation. **Gustavo Cernera:** Methodology, Formal analysis, Data curation. **Federica Zarrilli:** Formal analysis, Data curation. **Giuseppe Castaldo:** Writing – review & editing,

Conceptualization. **Felice Amato:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA - PubMed, Science (New York, N.Y.) 245(4922) (09/08/1989).
[2] DNA sequencing with chain-terminating inhibitors - PubMed, Proc. Natl. Acad. Sci. U. S. A. 74(12) (1977 Dec).
[3] Next-generation DNA sequencing - PubMed, Nat. Biotechnol. 26(10) (2008 Oct).
[4] Detection of two Alu insertions in the CFTR gene - PubMed, J. Cystic Fibrosis : Off. J. Eur. Cystic Fibrosis Soc. 7(1) (2008 Jan).
[5] Mobile elements and mammalian genome evolution - PubMed, Curr. Opin. Genet. Dev. 13(6) (2003 Dec).
[6] Hot L1s account for the bulk of retrotransposition in the human population - PubMed, Proc. Natl. Acad. Sci. U. S. A. 100(9) (04/29/2003).
[7] Alu repeats and human genomic diversity - PubMed, Nature reviews. Genetics 3(5) (2002 May).
[8] Increased Alu RNA processing in Alzheimer brains is linked to gene expression changes - PubMed, EMBO Rep. 22(5) (05/05/2021).
[9] The Alu neurodegeneration hypothesis: A primate-specific mechanism for neuronal transcription noise, mitochondrial dysfunction, and manifestation of neurodegenerative disease - PubMed, Alzheimer's & Dementia J. Alzheimer's Assoc. 13(7) (2017 Jul).
[10] Genomic investigation of alpha-synuclein multiplication and parkinsonism - PubMed, Ann. Neurol. 63(6) (2008 Jun).
[11] An Alu insert as the cause of a severe form of hemophilia A - PubMed, Acta Haematol. 106(3) (2001).
[12] A de novo Alu insertion results in neurofibromatosis type 1 - PubMed, Nature 353 (6347) (10/31/1991).
[13] Cystic Fibrosis-Related Diabetes (CFRD): Overview of Associated Genetic Factors., Diagnostics (Basel, Switzerland) 11(3) (2021/03/22).
[14] Genetic diseases that predispose to early liver cirrhosis - PubMed, Int. J. Hepatol. 2014 (2014).
[15] Molecular Diagnosis and Genetic Counseling of Cystic Fibrosis and Related Disorders: New Challenges - PubMed, Genes 11(6) (06/04/2020).
[16] Designing a simple multiplex ligation-dependent probe amplification (MLPA) assay for rapid detection of copy number variants in the genome - PubMed, Journal of genetics and genomics = Yi chuan xue bao 36(4) (2009 Apr).
[17] Membrane protein misassembly in disease - PubMed, Biochimica et biophysica acta 1818(4) (2012 Apr).
[18] ERAD: the long road to destruction - PubMed, Nat. Cell Biol. 7(8) (2005 Aug).
[19] Marginally hydrophobic transmembrane α-helices shaping membrane protein folding - PubMed, Protein Sci. Publ. Protein Soc. 24(7) (2015 Jul).
[20] How translocons select transmembrane helices - PubMed, Ann. Rev. Biophys. 37(1) (2008).
[21] Prediction of human mRNA donor and acceptor sites from the DNA sequence - PubMed, J. Mol. Biol. 220(1) (07/05/1991).
[22] Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information - PubMed, Nucl. Acids Res. 24(17) (09/01/1996).
[23] Ex vivo model predicted in vivo efficacy of CFTR modulator therapy in a child with rare genotype - PubMed, Mol. Genet. Genom. Med. 9(4) (2021 Apr).
[24] Two CFTR mutations within codon 970 differently impact on the chloride channel functionality - PubMed, Human Mutation 40(6) (2019 Jun).
[25] In silico analysis and theratyping of an ultra-rare CFTR genotype (W57G/A234D) in primary human rectal and nasal epithelial cells - PubMed, iScience 26(11) (10/12/2023).
[26] L1077P CFTR pathogenic variant function rescue by Elexacaftor-Tezacaftor-Ivacaftor in cystic fibrosis patient-derived air-liquid interface (ALI) cultures and organoids: in vitro guided personalized therapy of non-F508del patients - PubMed, Respiratory Res. 24(1) (09/06/2023).
[27] Extensive molecular analysis of patients bearing CFTR-related disorders - PubMed, J. Mol. Diagnos. JMD 14(1) (2012 Jan).
[28] A novel DHPLC-based procedure for the analysis of COL1A1 and COL1A2 mutations in osteogenesis imperfecta - PubMed, J. Mol. Diagnos. JMD 13(6) (2011 Nov).