Molecular Biology Reports (2020) 47:6463–6469
https://doi.org/10.1007/s11033-020-05677-z

SHORT COMMUNICATION

Novel frameshift variant of the *CFTR* gene: S511Lfs*2 from phenotype to molecular predictions

Thaiane Rispoli[1,2] · Grazielle Motta Rodrigues[2,3] · Mayara Jorgens Prado[1,2] · Leonardo Araújo Pinto[4] · Marcelo Tadday Rodrigues[5] · Cynthia Rocha Dullius[5] · Tarciana Grandi[2] · Cláudia Maria Dornelles da Silva[2] · José Eduardo Vargas[6] · Maurício Menegatti Rigo[7] · Maria Lucia Rossetti[1,8]

Received: 27 March 2020 / Revised: 9 July 2020 / Accepted: 19 July 2020 / Published online: 31 July 2020

Abstract

Cystic fibrosis (CF) is a genetic disease caused by variants in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. There are over 2,000 different pathogenic and non-pathogenic variants described in association with a broad clinical heterogeneity. In this work, we identified a novel variant S511Lfs*2 in *CFTR* gene that has not been reported in patients with CF. The patient was a female genotyped with c.1000C>T (legacy name: R334W) variant (pathogenic, CF-causing) and the novel variant (S511Lfs*2). We verified the amino acid sequence, the protein structure, and predicted the pathogenicity employing computational analysis. Our findings showed that S511Lfs*2 is a frameshift variant and suggest that it is associated with severe CF phenotype, as it leads to a lack of CFTR protein synthesis, and consequently the loss of its functional activity.

Keywords Cystic fibrosis · S511Lfs*2 variant · *CFTR* gene · CFTR protein · Frameshift variant

Introduction

Cystic fibrosis (CF-OMIM 219700) is an autosomal recessive inherited disorder caused by variants in the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*) gene [1]. The *CFTR* gene encodes a protein of 1480 residues in length (UniProt ID: P13569). This protein belongs to the ATP-binding cassette (ABC) transporter superfamily and its function is related to the transport of chloride ions across cell membranes [2, 3], regulation of airway fluid homeostasis [4–6], and defense against pathogens by the regulation of pH and ion content in the airway surface fluid layer [4, 6, 7].

The three-dimensional structure of CFTR protein is composed by five domains. Two transmembrane domains (TMDs), responsible for the chloride ion translocation pathway, two cytoplasmic Nucleotide Binding Domains (NBDs) with ATPase activity that controls channel gating, and a regulatory (R) domain that mediates CFTR activity upon phosphorylation [8]. More than 2000 variants have been already described in the *CFTR* gene (CFTR1, available in http://

Thaiane Rispoli and Grazielle Motta Rodrigues have contributed equally to the work.

✉ Thaiane Rispoli
thaianerispoli@gmail.com

1 Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

2 Centro de Desenvolvimento Científico e Tecnológico, Secretaria da Saúde do Estado do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

3 Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

4 Hospital São Lucas da PUCRS, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

5 Serviço de Fibrose Cística - Adultos, Serviço de Pneumologia, Hospital São Lucas da PUCRS, Porto Alegre, Rio Grande do Sul, Brazil

6 Instituto de Ciências Biológicas, Universidade de Passo Fundo, Passo Fundo, Rio Grande do Sul, Brazil

7 Programa de Pós-Graduação em Pediatria e Saúde da Criança, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

8 Programa de Pós-Graduação em Biologia Celular e Molecular Aplicada à Saúde, Universidade Luterana do Brazil, Canoas, Rio Grande do Sul, Brazil

www.genet.sickkids.on.ca/), and around 300 variants are considered pathogenic (CFTR2, available in https://www.cftr2.org/). These pathogenic variants are classified into major seven classes (IA, IB, II, III, IV, V, and VI), according to the molecular mechanisms' disruption: synthesis, traffic, or function of the protein [9, 10]. The classes IA, IB, II, and III are associated with a more severe prognosis due to defective protein production, processing, and regulation; the classes IV, V and VI are associated with milder prognosis, because CFTR protein maintains residual activity [9]. However, a mutation expanded classification can be considered with a combination of classes once a single mutation can cause defects in the CFTR protein by multiple mechanisms [11]. Therefore, the study of variants is fundamental to understand the patient prognostic and also to develop new treatments based on gene variants [12].

In this study, we described a novel frameshift variant (S511Lfs*2) resulted from the deletion of two thymines. The frameshift yielded a codon that specifies the amino acid leucine followed by a premature termination codon (PTC). We have combined the patient´s clinical features and structural bioinformatics data to propose the variant's effects in the CF phenotype.

Case report

Patient data collection

The subject from a CF Reference Center (Hospital São Lucas da PUCRS) was a 23-year-old female with a confirmed CF diagnosis at the age of twelve. She had the first sweat test intermediate (57 mmol/L) and the second abnormal (97 mmol/L). The patient had a history of respiratory infections with the airway colonized by *Staphylococcus aureus*, severe bronchiectasis—diagnosed by chest computed tomography (CT)—and mucoid impaction of the large airways. This subject also showed a mild airflow limitation with FVC = 2.75 (87% predicted), FEV1 = 2.12 (75% predicted) and a reduced FEV1/FVC ratio = 0.77 (86% predicted), recurrent pancreatitis, pancreatic insufficiency, and critical failure to thrive (weight Z-score of − 0.88 and height Z-score of − 2.04).

Molecular genetic analysis of the *CFTR* gene

Genomic DNA extraction from peripheral blood was performed by salting out assay [13]. The DNA was submitted to a molecular screening that analyses the presence of the eleven variants usually found in CF patients from south of Brazil and the main frequency variant c.1521_1523delCTT (legacy: F508del) found in CF patients worldwide. The eleven variants panel was carried out by SNaPshot [14] and the c.1521_1523delCTT was analysed by Sanger sequencing method as following. A PCR reaction was performed to amplify the exon 11 of the CFTR gene using specific primers (5′-TGAATCCTGAGCGTGATTTG-3′ and 5′-TGGGTAGTGTGAAGGGTTCAT-3′). The PCR product was purified with ExoSAP-IT PCR (Affymetrix, USA) according to the manufacture's protocol and used in the Sanger sequencing that was carried out with BigDye Terminator Cycle Sequencing kit v3.1 (Applied Biosystems, USA). Then the product was precipitated by ethanol/ethylenediaminetetraacetic acid (EDTA) protocol recommended by Applied Biosystems chemistry guide [15]. The sequencing system used was the Genetic Analyzer 3130xl (Applied Biosystems, USA). Sequences obtained were aligned with the CFTR gene sequence (e.g. RefSeq NG_016465.4) and examined using SeqMan software v.17 (DNAstar, USA).

Cloning of the novel variant

To precisely characterize the deletion, an allele-specific cloning and sequencing method was carried out. The exon 11 of the *CFTR* gene was amplified by PCR using specific primers that contains a target sequence at the 3′-end of the primer, a sequence for restriction enzyme (BamHI in the forward primer and EcoRI in the reverse) and 6pb at the 5′-end of the primer to ensure efficient DNA cleavage by the restriction enzymes (CFTR_BamHI_F 5′-tttttGGATCCTGAATCCTGAGCGTGATTTG-3′ and CFTR_EcoRI_R 5′-tttttGAATTGATGCTTTGATGACGCTTCTG-3′). The PCR product was precipitated using the sodium acetate/ethanol method [16]. Each 1 µg of DNA from PCR product purified and pUC18 vector were sequentially digested with EcoRI and BamHI (Thermo Fisher Scientific) according to the manufacturer's instructions. After that, the DNA reactions were precipitated with sodium acetate/ethanol method. Using a 1:3 molar ratio of vector:insert DNA (100 ng of cut plasmid and 26 ng of cut insert) and 3 units of T4 DNA Ligase (Promega) the ligation reaction was performed according to manufacturer's recommendations and incubated for 18 h at 22 °C. The DNA was precipitated with sodium acetate/ethanol method and then transformed into XL1-Blue competent cells that were plated on 25 mL of LB solid medium with 0.85 mg X-Gal, 1 mM IPTG and 2.5 mg ampicillin. The plate was incubated overnight at 37 °C. White colonies resulted from transformation were used as a template to perform a PCR with the M13 primers to confirm the correct ligation. The Sanger sequencing with M13 primers was performed as described above in the "Molecular genetic analysis of the *CFTR* gene" section.

Structure and pathogenicity prediction

The ExPASy software was used to obtain the translation of the nucleotide sequence to a protein sequence [17]. The putative protein structure was built using UCSF chimera software based on the structure found in Protein Data Bank (PDB ID: 5UAK) [18]. The MutationTaster software was used to predict the pathogenicity considering evolutionary conservation, mRNA, and protein structure/function through phyloP/phastCons scores [19]. The PhyloP measure the conservation or acceleration, either lineage-specific or across all branches and PhastCons the conservation scoring and identification of conserved elements [19].

Results

The SNaPshot genetic analysis identified a compound heterozygote. One variant was c.1000C>T (legacy: R334W), which has been previously reported in the CFTR2 database (Fig. 1). The other variant was a novel deletion variant found by Sanger sequencing; however, it was challenging to identify the deletion correctly (Fig. 2a). Thereby, to characterize this novel variant, we isolated the sequences by allele-specific cloning (Fig. 2b).

The novel variant, S511Lfs*2 was discovered in exon 11 of the *CFTR* gene, where two thymines at position g.98818_98819del (NG_016465.4) were deleted (Fig. 2b). As a result, this deletion formed the premature termination codon UGA after replacing serine for leucine being a frameshift variant type (Fig. 2c). The novel variant was deposited in Cystic Fibrosis Mutation Database (CFTR1) as well as in the public archive of human variations and phenotypes (ClinVar; SCV000987313) and in the public archive of Single Nucleotide Polymorphism database (dbSNP; rs1562898471).

In silico analysis in MutationTaster software suggests that this deletion can modify functions of key regions such as the ATP-binding domain and the PDZ-binding domain (Fig. 3). The value obtained of two conservation parameters used, PhyloP (1.528–4.75 value) and PhastCons (1 value), revealed that deleted thymines are in a conserved region in the *CFTR* gene corroborating with pathogenic features. Furthermore, this PTC promotes a loss of approximately 65% of full-length CFTR protein (Fig. 2c), possibly resulting in nonsense-mediated mRNA decay (NMD). The structure of the protein lacking amino acids after the termination codon is shown in Fig. 3.

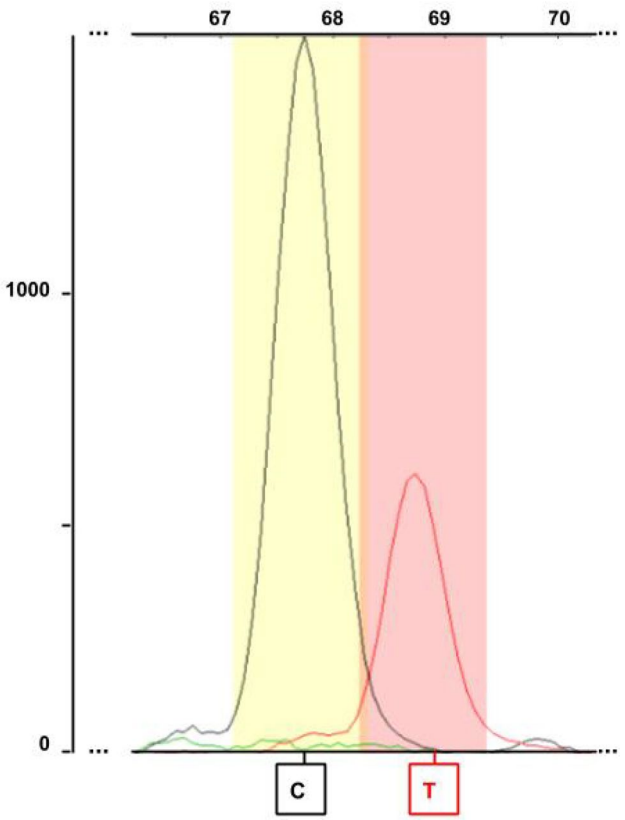

Fig. 1 Detection of c.1000C>T (legacy name: R334W) variant in heterozygosity in CFTR gene. Electropherogram of SNaPshot methodology: C allele (wild-type; black peak) and T allele (variant; red peak)

Discussion

This is the first study worldwide to describe the S511Lfs*2 variant identified in a Brazilian CF patient. Variants identification is essential for diagnosis purposes, specific treatment approaches, and genetic counseling [21]. We also demonstrate an in silico model to elucidate the protein alteration at a structural level.

The S511Lfs*2 variant leads to a change in the amino acids sequence (S511L) and promotes the appearance of a PTC. The PTC leads to the production of truncated and non-functional CFTR protein. It is possible that this modification may activate a post-transcriptional mechanism called nonsense-mediated mRNA decay (NMD) which detects and degrades the mutated transcript [22–24]. However, more studies would be required to evidence this mechanism. Thus, it is reasonable to classify this variant as a class IB due to the possible NMD activation, associated with pathogenicity prediction values, clinical data from the patient, and defective protein production. [9, 10]. It is important to emphasize that independently of the NMD activation, the protein structure would be strongly affected due to the absence of full-length functional CFTR protein (Figs. 2c and 3). Consequently, protein functionality would be lost.

Structural analysis of CFTR protein shows that the novel variant S511Lfs*2 is in the NBD1 domain, near the c.1521_1523delCTT (legacy: F508del) variant. The in

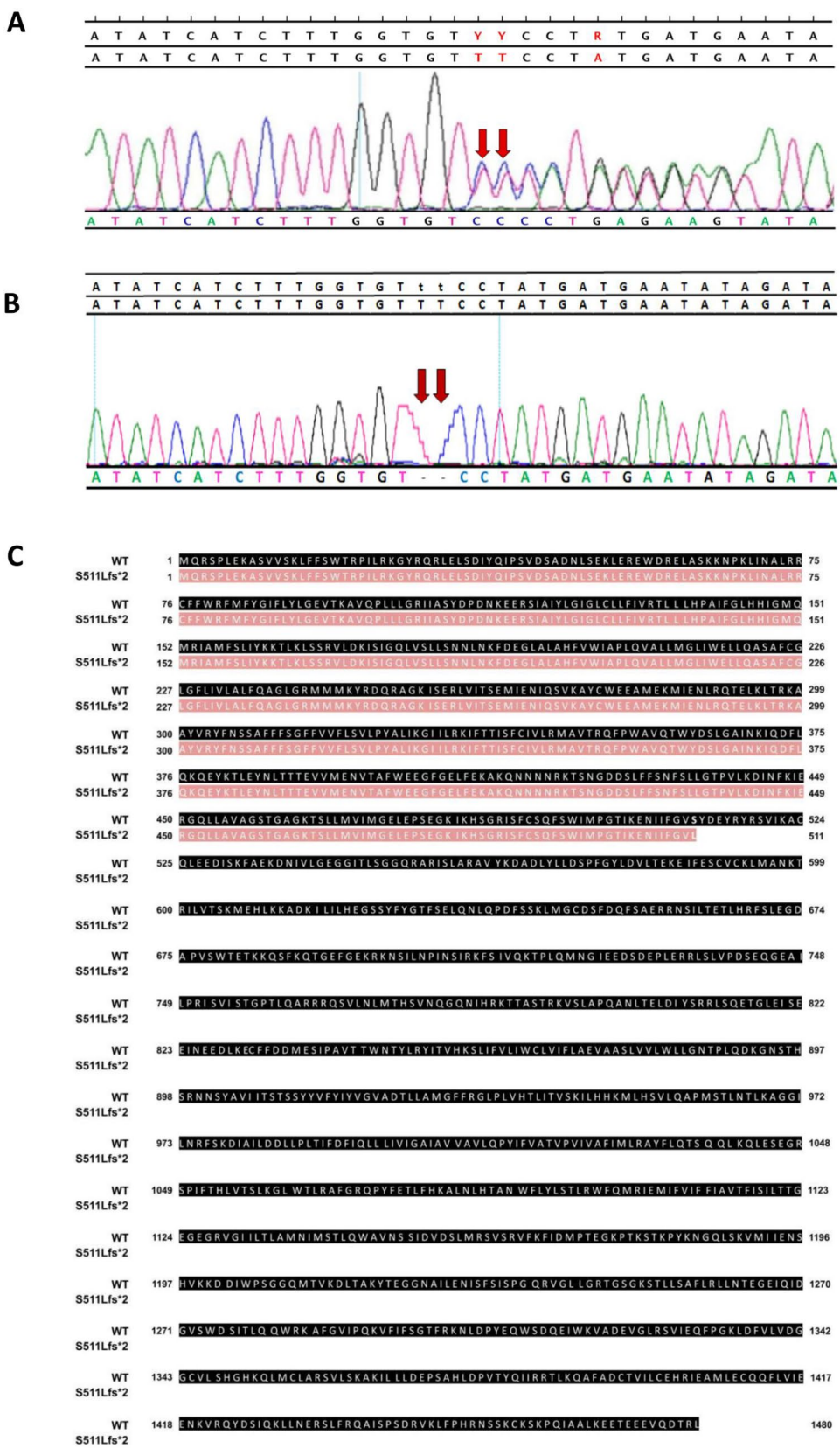

Fig. 2 From nucleotide sequence to the amino acid sequence of S511Lfs*2 protein. **a** Sequence electropherogram of the exon 11 (*CFTR* gene) showing the novel variant S511Lfs*2 in heterozygous state. **b** Detection of the novel S511Lfs*2 (deletion of two thymines) variant in one of the alleles after allele-specific cloning. The two red arrows confirm the deletion of two thymines in one allele. **c** CFTR protein amino acid sequence obtained from ExPASy software. The wild-type (black) and the mutated (pink) CFTR amino acid sequence are represented

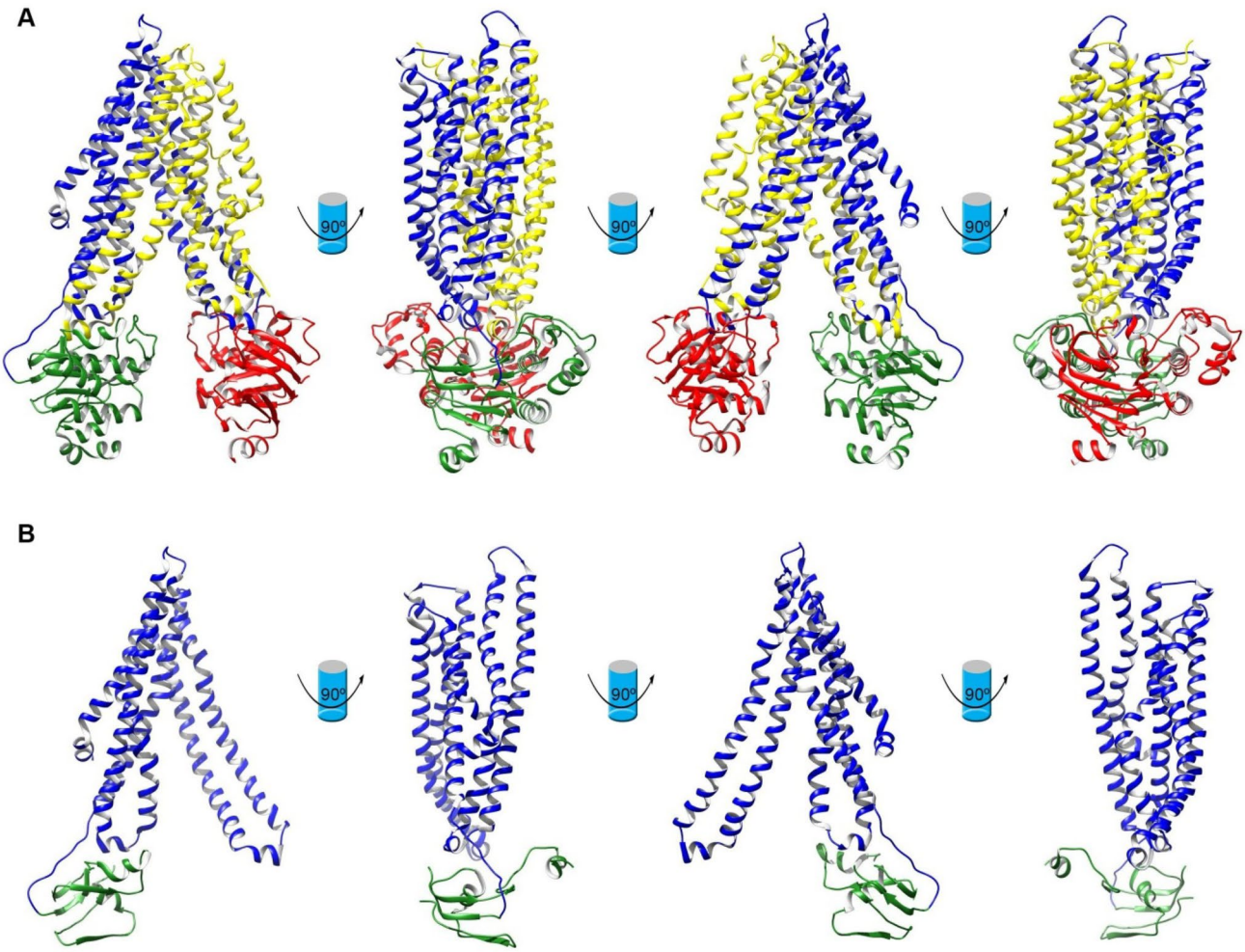

Fig. 3 CFTR three-dimensional structures. The dephosphorylated CFTR structure was modeled using the template 5UAK, from Protein DataBank [20]. The R domains are not represented. **a** Ribbon representation of the CFTR wild-type protein and **b** the CFTR S511Lfs*2 protein. The structure was rotated 90° to highlight all the faces of the protein. The functional domains TMD1 (blue), TMD2 (yellow), NBD1 (green), and NBD2 (red) are represented

silico analysis in this region showed a high conservation grade. Thus, any alterations in this domain have a deleterious impact on the protein by to affect essential regions, such as ATP-binding domain, which is essential to allow the open-closed status of the channel, and the PDZ-binding domain, that play a key role in anchoring proteins in the cell membrane [25].

Additional studies demonstrated that a single-residue deletion in the region 491–525 elicits an array of abnormalities in CFTR function [26]. Although it was found that specifically the amino acids V510 and S511 deletions did not cause problems in the processing of the CFTR protein, our data show that the new mutation S511Lfs*2 is caused by the deletion of two T nucleotides in different codons. This modification is a frameshift variant, causing the loss of the S511 amino acid and changing the reading phase, thus altering all subsequent positions within the critical region 511–525, in addition to also causing the formation of a premature codon stop.

Furthermore, this novel variant would potentially eliminate the R domain (Fig. 3), which is responsible by the channel activation and thus essential for the functionality of the protein. The loss of this domain was predicted based on sequence and structural analysis, since this domain was not resolved by the Electron Cryomicroscopy (cryo-EM) structure [20, 26].

We identified the S511Lfs*2 variant in compound heterozygosity with c.1000C>T (legacy: R334W), which represents the fourth most frequent disease-causing variant among Brazilian CF patients and leads to a mild CF phenotype (class IV) [10, 27]. The severity of the clinical features indicates that, although the S511Lfs*2 variant is associated with no CFTR protein synthesis, the c.1000C>T (legacy: R334W) variant confers the residual

CFTR channel activity, which could explain the delayed appearance of the symptoms and the late diagnosis of this patient [28, 29].

The knowledge of new variants, their classification, and mechanisms are essential to understand CF pathogenicity and drugs discovery able to correct defective proteins [10, 23]. Furthermore, in silico tools are a helpful strategy to predict the degree pathogenicity of the uncertain variants, as shown in this work.

Conclusions

In this study, Sanger sequencing detected a novel variant in a compound heterozygote that was confirmed by allele-specific analysis in vitro. The in silico analysis shows that S511Lfs*2 is a frameshift variant and generates a PTC which leads to the production of truncated and non-functional CFTR protein. Moreover, CF patients with this novel variant could benefit from therapies that improve the expression of full-length CFTR protein.

Acknowledgements The authors are grateful to Arnaldo Zaha for providing the laboratory facilities to carry out the allele-specific cloning assay of the novel variant.

Author contributions Thaiane Rispoli and Grazielle Rodrigues: designed the experiments, performed the experiments, analysed, and interpreted the data, drafted the manuscript, read, and approved the final manuscript. Mayara J. Prado: designed, drafted, and performed the in vitro allele-specific cloning, revised the manuscript, read, and approved the final manuscript. Leonardo Araújo Pinto, Marcelo Tadday Rodrigues, Cynthia Rocha Dullius: collected phenotypic data from patients, revised the manuscript, read, and approved the final manuscript. Tarciana Grandi, Cláudia Maria Dornelles da Silva, José Eduardo Vargas: revised the manuscript, read, and approved the final manuscript. Maurício Menegatti Rigo, Maria Lucia Rossetti: supervised the study, read, and approved the final manuscript.

Funding The study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES – Finance Code 001) and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS, 1260-2551/13-0).

Compliance with ethical standards

Conflict of interest Authors declared no conflict of interest.

Ethical approval The subject enrolled in the study was a volunteer from the CF Reference Center (Hospital São Lucas da PUCRS). The study was conducted according to the 1964 Helsinki declaration and all procedures involving human subjects were approved by the Local Ethics Committee of the Hospital São Lucas – PUCRS and Secretaria da Saúde do Estado do Rio Grande do Sul.

Informed consent Written informed consent was obtained from the subject enrolled in the study. Declaration of consent in the study from the participant is available from the correspondence author.

References

1. Ratjen F, Bell SC, Rowe SM, Goss CH, Quittner AL, Bush A (2015) Cystic fibrosis. Nat Rev Dis Prim 1:15010
2. Gadsby DC, Vergani P, Csanády L (2006) The ABC protein turned chloride channel whose failure causes cystic fibrosis. Nat Rev 440:477–483
3. Meng X, Wang Y, Wang X, Wrennall JA, Rimington TL, Li H, Cai Z, Ford RC, Sheppard DN (2017) Two small molecules restore stability to a subpopulation of the cystic fibrosis transmembrane conductance regulator with the predominant disease-causing mutation. J Biol Chem 292:3706–3719
4. Blouquit S, Regnier A, Dannhoffer L, Fermanian C, Naline E, Boucher R, Chinet T (2006) Ion and fluid transport properties of small airways in cystic fibrosis. Am J Respir Crit Care Med 174:299–305
5. Zhang L, Button B, Gabriel SE et al (2009) CFTR delivery to 25% of surface epithelial cells restores normal rates of mucus transport to human cystic fibrosis airway epithelium. PLoS Biol 7:e1000155
6. Shah VS, Meyerholz DK, Tang XX et al (2016) Airway acidification initiates host defense abnormalities in cystic fibrosis mice. Science 351:503–507
7. Coakley RD, Grubb BR, Paradiso AM, Gatzy JT, Johnson LG, Kreda SM, O'Neal WK, Boucher RC (2003) Abnormal surface liquid pH regulation by cultured cystic fibrosis bronchial epithelium. Proc Natl Acad Sci USA 100:16083–16088
8. Hwang T-C, Yeh J-T, Zhang J, Yu Y-C, Yeh H-I, Destefano S (2018) Structural mechanisms of CFTR function and dysfunction. J Gen Physiol 150:539–570
9. Marson FAL, Bertuzzo CS, Ribeiro JD (2016) Classification of CFTR mutation classes. Lancet Respir Med 4:e37–e38
10. De Boeck K, Amaral MD (2016) Progress in therapies for cystic fibrosis. Lancet Respir Med 4:662–674
11. Veit G, Avramescu RG, Chiang AN et al (2016) From CFTR biology toward combinatorial pharmacotherapy: expanded classification of cystic fibrosis mutations. Mol Biol Cell 27:424–433
12. Martins RS, Campos Junior M, dos Santos Moreira A, Marques Zembrzuski V, da Fonseca ACP, Abreu G, de Cabello M, de Cabello PH (2019) Identification of a novel large deletion and other copy number variations in the CFTR gene in patients with cystic fibrosis from a multiethnic population. Mol Genet Genomic Med 7:e645
13. Mousquer GT, Maciel LP, Pompeu Saraiva AC, Dalla Costa ER, Rosa Rossetti ML (2018) Validation of a quick and low-cost DNA extraction protocol applicable to long-stored blood samples. Anal Biochem 561–562:47–51
14. Rispoli T, Martins de Castro S, Grandi T, Prado M, Filippon L, Dornelles da Silva CM, Vargas JE, Rossetti LMR (2018) A low-cost and simple genetic screening for cystic fibrosis provided by the Brazilian Public Health System. J Pediatr 199:272-277.e3
15. Yang DL, Sauvageot R, Pentoney SL (2007) DNA sequencing by capillary electrophoresis. Electrophoresis. https://doi.org/10.1002/elps.200900218
16. Zeugen JA, Hartley JL (1985) Ethanol precipitation of DNA. Focus (Madison) 7:1–2
17. Artimo P, Jonnalagedda M, Arnold K et al (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40:W597–W603
18. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera: a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612
19. Schwarz JM, Cooper DN, Schuelke M, Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11:361–362

20. Liu F, Zhang Z, Csanády L, Gadsby DC, Chen J (2017) Molecular structure of the human CFTR Ion channel. Cell 169:85-95.e8
21. Patch C, Middleton A (2018) Genetic counselling in the era of genomic medicine. Br Med Bull 126:27–36
22. Lejeune F (2017) Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. BMB Rep 50:175–185
23. Zainal Abidin N, Haq IJ, Gardner AI, Brodlie M (2017) Ataluren in cystic fibrosis: development, clinical studies and where are we now? Expert Opin Pharmacother 18:1363–1371
24. Clarke LA, Awatade NT, Felício VM et al (2018) The effect of premature termination codon mutations on CFTR mRNA abundance in human nasal epithelium and intestinal organoids: a basis for read-through therapies in cystic fibrosis. Hum Mutat 40:23692
25. Guggino WB (2004) The cystic fibrosis transmembrane regulator forms macromolecular complexes with PDZ domain scaffold proteins. Proc Am Thorac Soc 1:28–32
26. Chen X, Zhu S, Zhenin M, Xu W, Bose SJ, Wong MPF, Leung GPH, Senderowitz H, Chen JH (2019) A defective flexible loop contributes to the processing and gating defects of the predominant cystic fibrosis-causing mutation. FASEB J 33:5126–5142
27. Grupo Brasileiro de Estudos de Fibrose Cística (2016) The Brazilian Cystic Fibrosis Patient Registry
28. Kristidis P, Bozon D, Corey J, Danuta Markiewicz M, Rommens J, Tsui L-C, Durie P (1992) Genetic determination of exocrine pancreatic function in cystic fibrosis. Am J Hum Genet 50:1178–1184
29. Sheppard DN, Rich DP, Ostedgaard LS, Gregory RJ, Smith AE, Welsh MJ (1993) Mutations in CFTR associated with mild-disease-form CI-channels with altered pore properties. Nature 362:160–164