



HAL
open science

CFTR gene variants, epidemiology and molecular pathology

C. Bareil, A. Bergounoux

► **To cite this version:**

C. Bareil, A. Bergounoux. CFTR gene variants, epidemiology and molecular pathology. Archives de Pédiatrie, 2020, 27, pp.eS8-eS12. 10.1016/S0929-693X(20)30044-0 . hal-03477908

HAL Id: hal-03477908

<https://hal.science/hal-03477908>

Submitted on 20 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CFTR gene variants, epidemiology and molecular pathology

C. Bareil^a, A. Bergougnoux^{a,b,*}

^aLaboratoire de Génétique Moléculaire, CHU de Montpellier, Montpellier, France

^bEA7402 Laboratoire de Génétique de Maladies Rares, Institut Universitaire de Recherche Clinique, Université de Montpellier, Montpellier, France.

ARTICLE INFO

Keywords:

CFTR
Variant
Mutational spectrum
Functional classification

ABSTRACT

Pathogenic variants of the *CFTR* gene are responsible for a broad phenotypic spectrum characterized by malfunction of some exocrine tissues, with an autosomal recessive mode of inheritance. More than 2,000 variants, distributed throughout the *CFTR* gene, have been identified, with different effects on the gene and protein expression and function. Genotype-phenotype correlation studies have associated severe variants with a typical multi-organ form of cystic fibrosis, while mild variants are involved in mono-symptomatic or adult-onset diseases, called CFTR-related disorders. However, the interpretation of rare variants remains challenging.

This review presents an overview of the epidemiology of *CFTR* variants worldwide and in France and describes the functional classification. Finally, some frequent cystic fibrosis-causing and mild *CFTR* variants are used as example to depict the molecular pathology of the *CFTR* locus. Finally, we give the recommendations concerning nomenclature and classification that are useful for appropriate genetic counseling.

© 2020 French Society of Pediatrics. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Cystic fibrosis (CF) is the most frequent life-restrictive genetic disorder in Caucasian populations with an autosomal recessive mode of inheritance. Registry-derived population estimates suggest that more than 72,000 people live with CF worldwide [1]. CF is caused by the presence of variants in both copies of the *Cystic Fibrosis Transmembrane conductance Regulator (CFTR)* gene, each inherited from one parent. *CFTR* pathogenic variants affect CFTR protein at different levels: synthesis (mRNA and protein), maturation, trafficking, and channel activity. The absence or dysfunction of the CFTR protein leads to ion flux perturbations in the epithelial cells of various organs.

More than 2,000 *CFTR* variants distributed throughout the gene and of varied clinical impacts have been identified, and each patient's phenotype is the result of the combination of two (or more) variants. CF-causing variants lead to severe CF clinical manifestations when combined. Moderate or mild variants, are involved in single-organ disorders called CFTR-Related Disorders (CFTR-RD: diffuse bronchiectasis, pancreatitis, male infertility with congenital bilateral absence of vas deferens). A small number of *CFTR* variants are associated with a broad phenotypic spectrum: from CF to CFTR-RD, or from CFTR-RD to no symptoms. Finally, there is a huge number of variants

with unknown clinical spectrum because they are very rare and have not been functionally studied yet (for more details, see clinical classification [2]).

In this review, we will describe the epidemiology of *CFTR* variants worldwide and in France. We will review the historical functional classification that was proposed 25 years ago [3]. Through examples of some frequent CF-causing and mild *CFTR* variants, we will describe key points (nomenclature, classification) concerning the molecular pathology of the *CFTR* locus that are useful for clinical practice to provide appropriate genetic counseling and consider personalized therapy.

2. Epidemiology of *CFTR* variants

The high number of *CFTR* variants and their worldwide distribution have been extensively incremented by the network of CF laboratories [4,5] since the gene discovery in 1989 [6,7].

Loss-of-function[#] *CFTR* variants have been identified in all 27 coding exons and their flanking intronic boundaries, and also in the promoter[#] and even in deep intronic regions (see Glossary in text box for terms followed by[#]; see also Supplementary Fig. 1A). Such DNA changes are mainly small substitutions or indels[#] of a few base pairs (98%) [8]. Large rearrangements that involve one or more exons represent 1-2% of all variants. This is very low compared

* Corresponding author.

E-mail address: anne.bergougnoux@inserm.fr (A. Bergougnoux).

with about 70% in the *DMD* gene involved in Duchenne muscular dystrophy, the most frequent X-linked recessive disorder.

The frequency of a given *CFTR* variant considerably varies depending on ethnic/geographic origins. For example, the p.Phe508del variant, that is the most frequent worldwide, represents 81% of CF-causing alleles[#] in some areas of Brittany and 24.5% in Turkey, with a clear northwest to southeast gradient among European populations [4]. Currently, the 29 most frequent variants in newborns with CF in France (included in the newborn screening programme panel) represent 87.1% of all CF alleles in this population (supplementary Table 1) [9], (for details see the article by M. P. Audrézet and A. Munck in this issue, “Newborn screening for CF in France: An exemplary national experience”). Other CF variants should be specifically searched in specific populations for optimal screening: for example, the deletion[#] that removes exons 2 and 3 frequent in Slavic populations, and the c.273+1G>A variation in some Jewish populations ([4], legacy name 405+1G>A). Most other variants are observed only in a small number of individuals. The list of novel variants is continuously increasing.

3. Consequences of variants according to their location in the *CFTR* gene

3.1. Exonic *CFTR* variants

Exon sequence changes represent 68% of the reported *CFTR* gene variants. Nucleotide substitutions in *CFTR* exons may result in missense[#] (39%), nonsense[#] (8%), splicing[#], or synonymous[#] variants (for a glossary see text box *Types of variants and their consequences*). Indels that alter the open-reading frame (i.e., frameshift) leading to a premature stop codon (PTC[#]), represent about 16% of distinct variants whereas in frame indels represent 2%. The pathogenicity of missense variants depends on the properties of the affected amino acids and their location in the *CFTR* channel sequence. For example, the p.Asn1303Lys (N1303K) variant is located in the NBD2 domain involved in *CFTR* protein folding, and destabilizes this domain. The p.Gly551Asp (G551D) variant, located in the consensus ATPase binding site, alters the *CFTR* channel function at the membrane (Supplementary Fig. 1A, 1B et 1C). Both variants have a severe effect and cause classical CF. Missense variants may also have a moderate effect or be neutral. Importantly, exonic variants, even if synonymous, may alter the normal mRNA splicing. For example, the c.3700A>G variant was initially predicted to be a missense variant and therefore named p.Ile1234Val. However, it is a splicing variant leading to an abnormal transcript that is not translated into a functional *CFTR* protein [10].

3.2. Intronic *CFTR* variants

Intronic changes located near exon boundaries, specifically at canonical splicing sites (i.e., acceptor AG (-2;-1), or donor GT (+1;+2), Supplementary Fig. 1A), typically alter splicing and lead to abnormal transcripts that are not translated into a functional protein. One of the most frequent consequences is exon skipping whereby one exon is excised and thus absent from the mRNA. For example, total exon skipping is observed for the frequent c.1585-1G>A (1717-1G>A) variant and partial exon skipping (coexistence of normal and mutated transcripts) is associated with the c.2657+5G>A (2789+5G>A) variant.

Overall, wherever their location (exonic, intronic or deep-intronic [11,12]), splicing variants represent almost 11% of all *CFTR* variants (CFMDB international database [13]). Their pathogenicity depends on the residual quantity of normal *CFTR* transcripts (see section 4.).

3.3. Large rearrangements in the *CFTR* gene

Large deletions or duplications that affect one or more *CFTR* exons are reported in 1-2% of all patients with CF worldwide. Except

some frequent rearrangements in some populations [4], most of them are limited to a small number of individuals. As all *CFTR* protein domains are needed for a functional channel, in most cases, these DNA rearrangements are associated with a CF phenotype [8]. Nevertheless, it has been suggested that in frame duplications[#] could lead to a less severe phenotype [14].

3.4. Variants in non coding *CFTR* regions

DNA changes upstream or downstream of the *CFTR* coding sequence that affect the *CFTR* mRNA quantity could slightly modify the clinical phenotype [15,16]. However, as their clinical impact is difficult to prove and no CF-causing variant has been identified in these distant regions, these variants are not currently explored in the routine diagnostic work-up.

3.5. Biological impact of *CFTR* variants: evidences and cautions

The biological consequences of the rarest *CFTR* variants are only presumed because of insufficient clinical, epidemiological, or functional data. The variant location and type may be suggestive of a given class in the functional classification (proposed by [3]) (see section 4). Some variants are classified as CF-causing because of their presumed severe consequence, such as nonsense, frameshift or splicing variants (Supplementary Table 1, text box *Demystifying the Variant Nomenclature*). Some frequent variants are easily considered as not pathogenic because of their high frequency in the general, healthy population. However, most nucleotide substitutions are rare missense variants for which the actual molecular impact and the associated clinical spectrum are difficult to predict. Moreover, the existence of complex alleles that combine two or three variants on the same parental allele complicate their interpretation and classification [17].

Comprehensive information provided by locus-specific databases (CFMDB [13], CFTR2 [18], CFTR-France [19] and ABCMdb [20]), the functional classification (see section 4) and clinical classification [2] greatly help clinicians and geneticists to offer adapted genetic counseling, to predict the clinical outcome, and to consider personalized therapy (for further details on databases, see the article by C. Raynal and H. Corvol in this issue, “Variant classifications, databases and genotype-phenotype correlations”).

4. Functional classification of *CFTR* variants

More than 25 years ago, Welsh and Smith proposed a functional classification of *CFTR* variants (class I to V) based on their consequences on *CFTR* protein function evaluated using *in vitro* chloride transport and cell function assays [3]. Figure 1 illustrates, with examples, the classes I to VI (class VI has been more recently proposed [21]) of *CFTR* pathogenic variants [22].

4.1. Interpretation of variant pathogenicity and functional classification: strengths and weaknesses

The concomitant presence of class I, II or III variants in the two *CFTR* gene copies of a patient is often associated with a severe phenotype, especially with pancreatic insufficiency. Conversely, the concomitant presence of class IV, V or VI variants, which lead to the production of a protein with some residual function, or with severe variants, is often associated with a mild phenotype, such as pancreatic sufficiency (for more details on the clinical classification [2], see the article by C. Raynal and H. Corvol in this issue). Although this functional classification appears to be an easy way to categorize variants for physicians and patients and to propose targeted therapies, there are still significant limitations.

The first and major limitation is that this classification is too partitioned. Indeed, many *CFTR* variants impair more than one single cellular process. For instance, the common p.Phe508del variant classically belongs to class II (absence of functional channels).

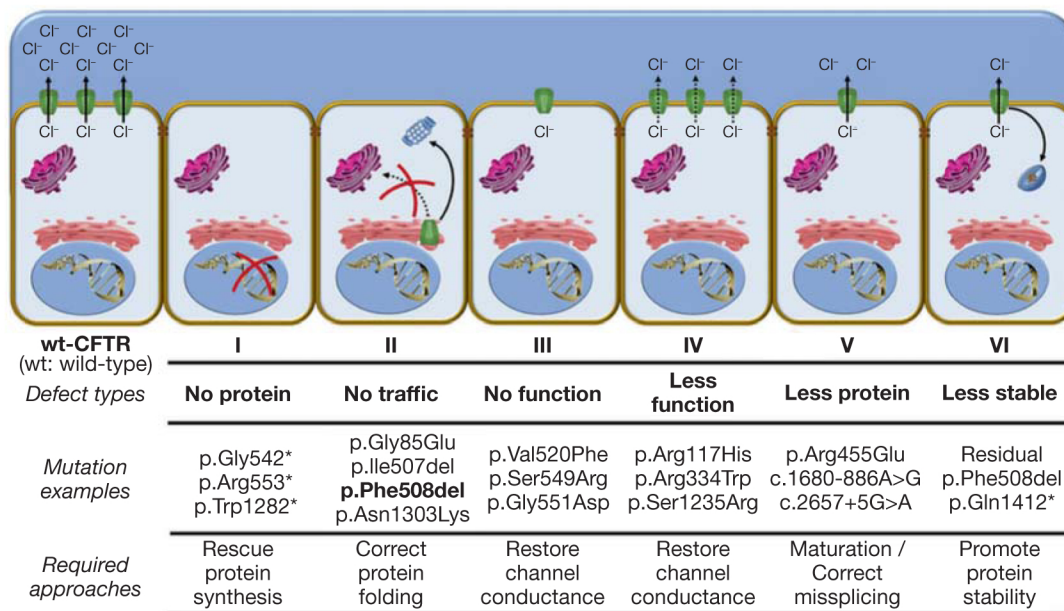


Fig. 1. Functional classification of *CFTR* variants and strategies for personalized therapies, adapted from [22].

Class I includes variants that lead to defective protein synthesis and absence of functional CFTR at the epithelial cell surface such as nonsense[#], splice sites[#] variants, insertions[#] or deletions[#] that result in a PTC[#].

Class II includes variants leading to defective CFTR protein processing and maturation (folding or incomplete glycosylation) and absence of functional CFTR channels at the epithelial cell surface, because of rapid proteasomal degradation. It includes numerous missense[#] variants and p.Phe508del, the most frequent variant worldwide.

Class III includes variants leading to defective channel regulation or “gating”. The CFTR protein is present at the apical membrane, but is not functional. Missense[#] variants, mostly located in the sequences encoding the nucleotide binding domains, do not impact the translation; the mutant proteins are properly processed, and addressed to the membrane. However, anion conductance and resistance to activation is markedly reduced thus preventing channel opening. The most widely studied is p.Gly551Asp (G551D), sensitive to potentiator ivacaftor, the first CFTR modulator approved worldwide.

Class IV includes variants leading to defective channel conductance and to the presence at the apical membrane of a CFTR protein with reduced anion conductance. Some missense[#] variants, mostly located in sequences encoding the transmembrane domains involved in the formation of the channel pore, result in the production of a CFTR protein with altered conductance properties, such as ion selectivity. p.Arg117His is a frequent variant of Class IV.

Class V includes variants that cause abnormal splicing[#] and that result in a mixture of correctly and aberrantly spliced *CFTR* mRNA, thus leading to reduced amount of normal CFTR protein. This class includes intronic variants, such as c.1210-12T[5] (“the 5T allele”, for details see “Mild variants on *CFTR* gene” on supplemental data), c.1680-886A>G (1811+1.6kbA>G) a recurrent deep-intronic CF variant, and some pathogenic variants of the promoter[#].

Class VI includes variants leading to a CFTR protein with defective stability and absence or severe reduction of CFTR at the cell surface. It includes some PTC[#] variants, such as p.Ser1455^{*}, producing a CFTR protein with residual function, thus leading in milder CF.

However, it also creates a gating defect (class III) observed in the few misfolded mutant proteins that reach the apical membrane. It also leads to a dramatic decrease in exocytic trafficking and peripheral stability due to impaired glycosylation of the mutant protein (class VI) [23]. A comprehensive classification would require a multiple analysis approach to investigate all possible effects of each variant.

The second limitation is that this functional classification does not take into account the other numerous regulatory functions (demonstrated or suggested) of the CFTR protein, such as bicarbonate channel and pH regulator [24].

Third, this classification remains incomplete. Most of the rare variants are currently not classified or may be misclassified, because 1) all consequences of a variant are not systematically studied and 2) *in vitro* experiments might not fully reflect *in vivo* conditions.

In the future, new models, such as intestinal organoids or airway primary cells, optimized for drug screening in function of the patient genetic background, could bring precious insights into the molecular defect of each variant and improve their classification [25] (for details, see the article by T. Bienvenu and T. Nguyen-Khoa in this issue, “Current and future diagnosis of cystic fibrosis: Performance and limitations”).

4.2. A new classification towards a therapeutic point of view

A modification of the functional classification has been recently suggested, based on a therapeutic approach rather than on the molecular and cellular defects [26-28], especially after

the launch of CFTR potentiators and correctors to improve CFTR function [29]. In addition, for more readability and for facilitating recruitment in clinical trials, a simplified two-classes categorisation is currently discussed. The first class would include variants leading to a very impaired or minimal CFTR function, such as class I variants of the functional classification (e.g., c.1585-1G>A). The second class would include variants leading to a residual CFTR function. Some variants, among both classes, could lead to a CFTR protein amenable to treatment with CFTR modulators such as potentiators (e.g., p.Gly551Asp, minimal function, or p.Arg117His, residual function) or an association of one or two correctors and one potentiator (e.g., p.Phe508del, minimal function). (For more details, see the article by I. Fajac and E. Girodon in this issue, “Genomically-guided therapies: A new era for cystic fibrosis”).

5. Mild variants associated with CFTR-RD phenotypes

The main challenge in genetic counseling for CF is to discriminate among variants according to the severity of their impact. Indeed, only CF-causing variants are considered for carrier screening and prenatal diagnosis. Moreover, some *CFTR* variants, classified as mild or considered not fully penetrant[#] (from CFTR-RD to no symptom), are quite frequent (see supplementary data). Clinicians should be aware of them [30] because they largely contribute to the phenotypic spectrum of patients with CFTR-RD (see the article by A. Pagin et al. in this issue, “Genetic diagnosis in practice: From cystic fibrosis to CFTR-related disorders”). It is accepted that

5-10% of normal, full-length *CFTR* transcripts should be sufficient to prevent a severe CF form. Nevertheless, a reduced quantity of transcripts (10-30%) is compatible with a CFTR-RD phenotype [31].

Conclusion

Since the discovery of the *CFTR* gene 30 years ago, the number of variants reached a total number of 2,100 in 2019. A set of frequent CF-causing variants, the frequencies of which vary depending on the geographic and ethnic origin, are currently assessed for first-tier screening (e.g., for neonates in national screening programmes). The second-tier *CFTR* analysis that covers all coding *CFTR* regions constantly reveals new variants with a hard to predict phenotype, especially for missense variants.

The current challenges concern the interpretation of rare variants, for which epidemiological data are missing. Functional *in vitro* studies, if combined with approaches to evaluate several defects (missense, splicing), would offer, despite technical and cost limitations, a comprehensive way to cluster variants in an effective functional classification, and will lead ultimately to strategies for class-specific therapeutic approaches.

In the future, a complementary strategy, based on the positive response to potentiators and/or correctors in *ex vivo* assays that take into account the patient genetic background (e.g., organoids) will allow us to propose a personalized therapy for patients with CF and CFTR-RD.

Take-home message

- The terms "variant", "mutation", "variation", "DNA change" or "DNA alteration" are synonymous and do not infer any clinical impact. According to the international recommendations, all DNA changes should be called "variant", irrespective of their consequences.
- There is a huge molecular heterogeneity in CF- and CFTR-RD-causing variants, with almost 2,100 variants, and the list is continuously increasing.
- The frequency of *CFTR* gene variants varies depending on the ethnic/geographical origins. Thus, the screening test has to be adapted to offer optimal coverage for specific populations.
- Frequent CF-causing variants are included in commercial screening kits, whereas the detection of rare variants requires the complete scanning of *CFTR* coding regions and their intronic boundaries, but also the specific search of recurrent deep-intronic variants.
- The impact of *CFTR* variants hugely varies. Assessment, especially of missense variants, remains challenging because of i) the existence of neutral or mild variants, ii) the existence of complex alleles, and iii) the potential association with a large phenotypic spectrum.
- Different variant classifications are in use to help geneticists, scientists and physicians to interpret the impact of variants:
 - A functional classification, based on the assessment of the variant molecular impact on the CFTR protein (class I to VI)
 - A clinical classification, based on the phenotype of patients carrying the variant and a CF-causing variant in *trans*: CF-causing, CFTR-RD-causing, variable clinical consequence (associated with CF and CFTR-RD), unknown or uncertain clinical significance, neutral or non-disease-causing.

Demystifying the Variant Nomenclature

The Human Genome Variation Society (HGVS) sets standards to describe genetic changes in all genes to provide a universal and unequivocal variant name (<https://varnomen.hgvs.org/>). The only unequivocal and accurate description of a genetic variant is its description at the DNA level. However, for *CFTR* variants, two other nomenclatures are used: the description at the protein level, and the legacy or traditional name (examples in supplementary Table 1).

A simple example: c.3909C>G, p.Asn1303Lys, N1303K

– Description at the DNA level: **c.3909C>G** is a Cytosine to Guanine nucleotide substitution at position No. 3909 of the *CFTR* complementary DNA (cDNA sequence). For patient diagnosis, the reference sequence is most often the "coding DNA sequence". "c." indicates that the variant is referenced according to the nucleotide position within the *CFTR* coding sequence, corresponding to the messenger RNA (Supplementary Fig. 1B). Variants in gene promoters are mentioned as "c.-" with the number of nucleotides before the +1 translation# start site.

– Description at the protein level: **p.Asn1303Lys** predicts an Asparagine to Lysine amino acid substitution at position No. 1303 in the CFTR protein. "p." indicates that the variant is referenced according to the amino acid position within the CFTR protein sequence (Supplementary Fig. 1C). The protein name is usually inferred without experimental proof, and therefore often corresponds to the predicted amino acid change rather than to the result of the DNA variant.

– Legacy name: **N1303K**, equivalent to p.Asn1303Lys, according to the single-letter amino acid code. This nomenclature is still found because it has been used in the literature for 30 years and because the HGVS recommendations are still evolving.

Glossary (terms with # in the text)

Allele: One of several alternative forms of a gene.

Loss-of-function variant: A type of variant in which the altered gene product lacks the molecular function of the wild type gene.

Penetrance: The penetrance of a disease for a disease-causing variant or genotype is the proportion of individuals with the variation or genotype who exhibit clinical symptoms.

Promoter: DNA regulatory region that leads to initiation of gene transcription#; it is located upstream and near the transcription start sites (TSS) of the gene.

Transcription: Synthesis of RNA from DNA.

Transcription factors: Proteins that control the transcription rate of genetic information from DNA to RNA, by binding to a specific DNA sequence.

Translation: Synthesis of an amino acid sequence (peptide or protein product) from the messenger RNA.

Types of variants and their consequences (including terms with # in the text)

Deletion: One or more nucleotides are missing.

Frameshift: An indel (insertion and/or deletion) in a DNA sequence of a number of nucleotides that is not divisible by three (as codons), and thus changes the reading frame.

In frame deletion/insertion: Opposite to frameshift, this DNA change does not modify the reading frame. Thus, one or more amino acids are missing/inserted inside the protein.

Indel: Insertion and/or deletion of one or more nucleotides

Insertion/duplication: One or more nucleotides are added.

Missense: A nucleotide substitution that predicts an amino acid substitution.

Nonsense: A nucleotide substitution that occurs in a codon specifying an amino acid and changes it to a premature stop codon (PTC). It results in a shorter than normal protein.

PTC: Premature stop codon in the mRNA sequence. It is usually caused by a nonsense variant or frameshift variant. PTC variants generally cause mRNA degradation by a cellular process called Nonsense-Mediated Decay (NMD) and the lack of protein production rather than truncated proteins.

Splicing: A genetic alteration in the DNA sequence that occurs at nearly any location in the gene and can disrupt RNA splicing, thus resulting in the complete or partial loss of exons or the inclusion of intronic sequences, hence an altered protein-coding sequence.

Synonymous: A change in the DNA sequence that does not change the encoded amino acid due to genetic code redundancy.

Disclosure of interest

The authors declare they have no conflict of interest.

This article is part of the eSupplement *Genetics of Cystic Fibrosis* published with institutional support from Vertex.

Acknowledgements

We thank Emmanuelle Girodon-Boulandet, Isabelle Fajac and Caroline Raynal for precious discussions.

References

- [1] Jackson AD, Goss CH. Epidemiology of CF: How registries can be used to advance our understanding of the CF population. *J Cyst Fibros* 2018;17(3):297-305.
- [2] Castellani C, Cuppens H, Macek M, et al. Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. *J Cyst Fibros* 2008;7(3):179-96.
- [3] Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* 1993;73(7):1251-4.
- [4] Bobadilla JL, Macek M, Fine JP, et al. Cystic fibrosis: a worldwide analysis of CFTR mutations--correlation with incidence data and application to screening. *Hum Mutat* 2002;19(6):575-606.
- [5] Claustres M, Guittard C, Bozon D, et al. Spectrum of CFTR mutations in cystic fibrosis and in congenital absence of the vas deferens in France. *Hum Mutat* 2000;16(2):143-56.
- [6] Rommens JM, Iannuzzi MC, Kerem B, et al. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* 1989;245(4922):1059-65.
- [7] Riordan JR, Rommens JM, Kerem B, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989;245(4922):1066-73.
- [8] Férec C, Casals T, Chuzhanova N, et al. Gross genomic rearrangements involving deletions in the CFTR gene: characterization of six new events from a large cohort of hitherto unidentified cystic fibrosis chromosomes and meta-analysis of the underlying mechanisms. *Eur J Hum Genet* 2006;14(5):567-76.
- [9] Audrézet MP, Munck A, Scotet V, et al. Comprehensive CFTR gene analysis of the French cystic fibrosis screened newborn cohort: implications for diagnosis, genetic counseling, and mutation-specific therapy. *Genet Med* 2015;17(2):108-16.
- [10] Ramalho AS, Clarke LA, Sousa M, et al. Comparative ex vivo, in vitro and in silico analyses of a CFTR splicing mutation: Importance of functional studies to establish disease liability of mutations. *J Cyst Fibros* 2016;15(1):21-33.
- [11] Bergougnot A, Délétang K, Pommier A, et al. Functional characterization and phenotypic spectrum of three recurrent disease-causing deep intronic variants of the CFTR gene. *J Cyst Fibros* 2019;18(4):468-75.
- [12] Chillón M, Dörk T, Casals T, et al. A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811+1.6kbA-->G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype. *Am J Hum Genet* 1995;56(3):623-9.
- [13] CFMDB: <http://www.genet.sickkids.on.ca/cftr/app>
- [14] Taulan M, Viart V, Theze C, et al. Identification of a novel duplication CFTRdup2 and functional impact of large rearrangements identified in the CFTR gene. *Gene* 2012;500(2):194-8.
- [15] Romey MC, Pallares-Ruiz N, Mange A, et al. A naturally occurring sequence variation that creates a YY1 element is associated with increased cystic fibrosis transmembrane conductance regulator gene expression. *J Biol Chem* 2000;275(5):3561-7.
- [16] Lopez E, Viart V, Guittard C, et al. Variants in CFTR untranslated regions are associated with congenital bilateral absence of the vas deferens. *J Med Genet* 2011;48(3):152-9.
- [17] Terlizzi V, Castaldo G, Salvatore D, et al. Genotype-phenotype correlation and functional studies in patients with cystic fibrosis bearing CFTR complex alleles. *J Med Genet* 2017;54(4):224-35.
- [18] CFTR2 database: <https://cftr2.org/>
- [19] CFTR-France database: <https://cftr.iurc.montp.inserm.fr/cftr/>
- [20] Gyimesi G, Borsodi D, Sarankó H, et al. ABCMdb: A database for the comparative analysis of protein mutations in ABC transporters, and a potential framework for a general application. <http://abcmdb.hegelab.org/>
- [21] Haardt M, Benharouga M, Lechardeur D, et al. C-terminal truncations destabilize the cystic fibrosis transmembrane conductance regulator without impairing its biogenesis. A novel class of mutation. *J Biol Chem* 1999;274(31):21873-7.
- [22] Lopes-Pacheco M. CFTR Modulators: Shedding Light on Precision Medicine for Cystic Fibrosis. *Front Pharmacol* 2016;7.
- [23] Veit G, Avramescu RG, Chiang AN, et al. From CFTR biology toward combinatorial pharmacotherapy: expanded classification of cystic fibrosis mutations. *Mol Biol Cell* 2016;27(3):424-33.
- [24] Saint-Criq V, Gray MA. Role of CFTR in epithelial physiology. *Cell Mol Life Sci* 2017;74(1):93-115.
- [25] Dekkers JF, Berkers G, Kruisselbrink E, et al. Characterizing responses to CFTR-modulating drugs using rectal organoids derived from subjects with cystic fibrosis. *Sci Transl Med* 2016;8(344):344ra84.
- [26] De Boeck K, Amaral MD. Progress in therapies for cystic fibrosis. *Lancet Respir Med* 2016;4(8):662-74.
- [27] Marson FAL, Bertuzzo CS, Ribeiro JD. Classification of CFTR mutation classes. *Lancet Respir Med* 2016;4(8):e37-8.
- [28] Stanke F, Tümmler B. Classification of CFTR mutation classes. *Lancet Respir Med* 2016;4(8):e36.
- [29] Han ST, Rab A, Pellicore MJ, et al. Residual function of cystic fibrosis mutants predicts response to small molecule CFTR modulators. *JCI Insight* 2018;3(14).
- [30] Boussarroue A, Bergougnot A, Raynal C, et al. Pitfalls in interpretation of CFTR variants in the context of incidental findings. *Hum Mutat* 2019;40(12):2239-46.
- [31] Amaral MD. Processing of CFTR: traversing the cellular maze--how much CFTR needs to go through to avoid cystic fibrosis? *Pediatr Pulmonol* 2005;39(6):479-91.