Author manuscript, published in "Journal of Molecular Diagnostics 2007;9(5):582-588"
DOI : 10.2353/jmoldx.2007.070040

inserm-00185153, version 1 - 10 Nov 2009

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

Corinne Bareil[1], Caroline Guittard[1], Jean-Pierre Altieri[1], Carine Templin[1], Mireille Claustres[1,2,3], Marie des Georges[1,3]

Address:

1- CHU Montpellier, Hôpital Arnaud de Villeneuve, Laboratoire de Génétique Moléculaire, Montpellier, F-34000 France;

2- Université MONTPELLIER1, UFR Médecine, Laboratoire de Génétique Moléculaire, Montpellier, F-34000 France;

3- Inserm, U 827, MONTPELLIER, F-34000 France;

Short running head: Improved Molecular Diagnosis of CBAVD

Corresponding author and reprint request: Marie des Georges

Laboratoire de Génétique Moléculaire

Institut Universitaire de Recherche Clinique

641 av du doyen Gaston Giraud

34093 MONTPELLIER CEDEX 5

FRANCE

Phone: + 33 467 41 53 60, Fax: + 33 467 41 53 65

marie.desgeorges@montp.inserm.fr

inserm-00185153, version 1 - 10 Nov 2009

Available commercial kits only screen for the most common cystic fibrosis transmembrane conductance regulator (*CFTR*) mutations causing classic cystic fibrosis and for the Tn variant in IVS8. However, full scanning of *CFTR* is needed for the diagnosis of patients with cystic fibrosis or *CFTR*-related disorders (including congenital bilateral absence of the vas deferens) bearing rare mutations. Standard strategies for detecting point mutations rely on extensive scanning of the gene by denaturing gradient gel electrophoresis or denaturing high performance liquid chromatography, which are time-consuming. Moreover the haplotyping of IVS8-(TG)m and Tn tracts is still challenging despite several recent improvements. We have optimized both the detection of mutations and the haplotyping of IVS8 polyvariants in developing two methods: i) a rapid and robust direct sequence analysis of all exons/flanking introns of the *CFTR* gene based on single condition touchdown amplification/sequencing in 96-well plates, ii) a fluorescent assay which allows haplotyping of IVS8-(TG)mTn even without family linkage study. Combined with search for rare large rearrangements, this strategy detected 87.9% of *CFTR* defects in congenital bilateral absence of the vas deferens patients, a proportion considerably higher than those usually reported.

These highly efficient tests, scanning each sample in a few days, greatly improve the genotyping of patients with *CFTR*-related symptoms and may be particularly important in emergency situations such as fetus with hyperechogenic bowel suggestive of cystic fibrosis.

inserm-00185153, version 1 - 10 Nov 2009

Introduction

Mutations in the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*) gene are responsible for cystic fibrosis (CF, MIM 219700) and isolated congenital bilateral absence of the vas deferens (CBAVD, MIM 277180). More than 96% of the 1,500 *CFTR* defects reported so far are point mutations altering a few bases or only one base (http://www.genet.sickkids.on.ca), while an unknown proportion of CFTR dysfunctions are caused by large genomic rearrangements such as large deletions [1-7]. In CBAVD, 88% of patients found with two *CFTR* mutations carry a severe mutation (no CFTR function) *in trans* to a mild mutation (residual function) and 12% carry two mild mutations [8]. The vast majority of mutations in CBAVD are not detected by routine panels, designed to test up to 30 common severe CF mutations; they are scattered over the whole gene, so that all the exons and their flanking introns have to be extensively scanned to reach acceptable rates of mutation detection [8, 9]. The laborious but powerful manual DGGE (*Denaturing Gradient Gel Electrophoresis*) or DHPLC (*Denaturing High Performance Liquid Chromatography*) techniques represented until recently the most useful approaches for mutation detection in CBAVD. However, these techniques are unable to determine the length variants localized at the polypyrimidine locus upstream to the splice acceptor site of intron 8 (polyTG followed by polyT repeats) which affect the splicing efficiency of exon 9 and act as genetic modifiers of CFTR function. Five variants, (TG)9 to (TG)13, are known in the (TG)m tract, whereas up to seven different alleles have been reported in the Tn tract (common alleles with 9, 7, or 5 thymidines and rare alleles with 3 [10], 6 [11], 10 [12] or 11 [13] thymidines. In Caucasian populations, the frequency of IVS8-T5 allele in CBAVD patients (30%) is 6 times higher than in the general population (5%), and 34% of men with CBAVD have inherited a *CFTR* mutation on one gene and IVS8-T5 on the other one, making this combination the most common cause of CBAVD. The T5 variant is associated with high levels of exon 9 skipping

inserm-00185153, version 1 - 10 Nov 2009

which result in the production of a non-functional CFTR protein. However the splicing efficiency of the IVS8-T5 allele shows inter- and intra-individual variability, and its incomplete penetrance in CBAVD is largely influenced by the IVS8-(TG)m tract [14, 15]. The determination of (TG)m repeat number is predictive of pathogenic T5 alleles [16]. Longer TG repeats increase exon 9 skipping and raise the proportion of non functional CFTR protein, which emphasizes the importance of assessing the length not only of the Tn, but also of the (TG)m tract for diagnostic purposes. As the two repeats act in concert when modulating exon 9 inclusion or skipping in the CFTR mRNA, reliable haplotyping is an additional prerequisite for a meaningful genetic testing at this locus.

We have developed two strategies that improve both the identification of mutations and the haplotyping of IVS8 repeats. First, a method based on single condition touchdown amplification (SiCTA) in a 96-well plate format allows the rapid direct sequence analysis of all *CFTR* exons and flanking introns together with a portion of IVS11 and IVS19 sequences (which carry common mutations). Second, a fluorescent assay determines the length of the IVS8-(TG)m and Tn tracts of both alleles and their haplotypes even when familial segregation cannot be studied. These two simple, rapid and reliable assays can routinely be performed in a few working days.

inserm-00185153, version 1 - 10 Nov 2009

Materials and methods

Samples

We analyzed or re-analyzed a collection of 182 samples (previous *CFTR* analysis of 85 of these samples had been included as part of a collaborative study [8]). Clinical diagnosis of CBAVD was based on clinical examination with impalpable vas deferens, transrectal ultrasonography, semen analysis (volume, pH and sperm count in accordance with the World Health Organization guidelines – WHO, 1992), low concentrations of fructose and citrate. Patients with renal abnormalities were excluded. Informed consent was obtained from all patients. DNA was extracted from peripheral blood samples by using standard procedures.

Classic protocols for analysis of *CFTR* mutations and IVS8-Tn alleles

A complete scan of the 27 coding/flanking sequences of the *CFTR* gene was performed either by DGGE or by DHPLC. In addition, two intronic mutations, 1811+1.6kbA>G in IVS11 and 3849+10kbC>T in IVS19 and variations at locus IVS8-Tn were screened by specific PCR-restriction tests. Samples showing abnormal profiles were re-amplified from genomic DNA and directly sequenced with the BigDye Terminator v1.1 cycle sequencing kit from Applied Biosystems (Warrington, UK). Samples found with only one or no *CFTR* disease-causing mutation were further investigated for large rearrangements such as large deletions by a semi-quantitative fluorescent SQF-PCR assay previously developed in our laboratory that uses three multiplex PCRs covering the entire gene [17]. In most cases, the *cis* vs. *trans* status of the alterations was obtained by familial segregation.

New protocol for haplotyping IVS8-(TG)mTn repeats

A fluorescent assay based on three PCRs was designed. Exon 9 was first amplified using primers 9i5/9i3 described by Zielenski et *al*. [18] and purified according to the manufacturer's recommendations by QIAquick PCR purification kit (Qiagen , Hilden, Germany) in order to remove the unincorporated primers. An aliquot of this amplicon was then used as a template

inserm-00185153, version 1 - 10 Nov 2009

for three internal fluorescent PCRs amplifying the (TG)m and the Tn tracts separately, and the (TG)mTn tracts simultaneously (Table 1, Figure 1). Each amplification was carried out in a final volume of 100 µl containing PCR Buffer 10X, 20 mM of each dNTP, 10 pmol of each primer, 1U of Taq Polymerase (Applied Biosystems, Branchburg, NJ) and 1 µl of PCR1. After an initial denaturation step at 94°C for 2 minutes, 22 cycles were performed with denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec and extension at 72°C for 1 minute, followed by a final extension step at 72°C for 30 minutes. Aliquots of the 3 amplicons ((TG)m or Tn repeats, and (TG)mTn haplotypes) were pooled and separated by multicapillary electrophoresis on an ABI 3130xl Genetic Analyzer running GeneMapper v4.0 for allele identification. A control DNA with fully characterized IVS8-(TG)mTn haplotypes was added to the run.

New protocol for *CFTR* analysis: Single Condition Touchdown Amplification/sequencing

The 27 coding/flanking sequences and the portions of introns 11 and 19 that contain the common mutations 1811+1.6kbA>G and 3849+10kbC>T, respectively, were amplified in 30 separate amplicons obtained under single PCR cycling conditions in a 96-well plate and sequenced. For exons 6b and 9, two sequencing reactions differing at the 5'end were necessary to read both strands and avoid two polymorphic regions (IVS6a-(GATT)n and IVS8-(TG)mTn repeats, respectively). Due to its large size, exon 13 was analyzed by use of two overlapping amplicons. The sequences of the primers were derived from Zielenski *et al.*, [18] or Fanen *et al.*, [19] with the exception of IVS11 [20] and IVS19 [21]; reverse primers for exons 1, 2 and 14b, forward primers for exons 2, 17a and IVS11 had to be slightly modified to ensure specific amplification. For all studied regions excepted exons 6b and 9, PCRs and sequence reactions were performed using the same primers. The primer stocks for PCRs (a mixture of reverse and forward primers, Table 2) and for sequence reactions were

inserm-00185153, version 1 - 10 Nov 2009

stored at -20°C in 96-well plates so that both PCRs and sequence reactions set-up can be done with pipetting robot or multichannel pipettors. PCRs were performed in 96-well plates, in a 25-µl final volume containing 1X PCR Master Mix (Promega, Madison, WI), 3.2 pmol of each PCR primer and 10 ng of genomic DNA. The use of a touchdown PCR protocol [22] allows a single amplification condition for all the exons: a denaturation step at 94°C for 5 min, 10 touchdown cycles with annealing temperature decreasing 1°C per cycle (denaturation 94°C for 30 sec, annealing 60°C for 40 sec, primer extension 72°C for 1 min), 30 cycles at the final touchdown temperature (50°C) and a final extension step at 72°C for 8 min. To validate specific size and quantity of amplicons, 5 µl of the PCR products were checked by 2% agarose gel electrophoresis. Five µl of each amplicon were then transferred to a 96-well plate and treated with ExoSAP-IT (USB Corporation, Ohio, USA) for unused nucleotides and primers removal, according to the manufacturer's recommendations. Depending on the number of patients to be sequenced, a single 96-well plate can be used to analyze the *CFTR* gene either of one patient in both directions (64 sequences) or of three patients in one direction (96 sequences). Sequence reactions (5 µl final volume) were assembled by transfer of 1.6 pmol of sequencing primer in a 96-well plate along with a 3-µl mixture of 1 µl ABI Prism BigDye terminators (version 1.1), 1 µl 5X buffer mix and 1 µl of purified PCR product. Cycling conditions followed manufacturer's recommendations. Sequencing products were purified on a 96-well Montage™ SEQ96 Sequencing Reaction Cleanup Kit (Millipore, Billerica, MA). The samples were run on an ABI 3130xl Genetic Analyzer (Applied Biosystems). The resulting sequences (10,254 bases in total) were analyzed for mutations, independently by two reviewers, using ABI SeqScape v2.5 automated assembly and basecalling software. Setting of the basecaller was according to the manufacturer, and recorded any base with a second peak of >5% as mixed. All mutations were confirmed by DGGE, DHPLC or restriction analysis from a PCR using a different primer pair and a new

DNA dilution. For convenience to the clinicians, mutations and polymorphisms were named according to the nomenclature that is used by the members of the CF Genetic Analysis Consortium (http://www.genet.sickkids.on.ca).

inserm-00185153, version 1 - 10 Nov 2009

inserm-00185153, version 1 - 10 Nov 2009

Results

Classic protocols for analysis of the *CFTR* gene

The protocol applied to 182 CBAVD patients allowed the identification of 87 different mutations scattered over the minimal promoter, 23 exons, and 11 introns, including 4 complex alleles (p.[Arg74Trp;Val201Met;Asp1270Asn], p.[Asp443Tyr;Gly576Ala;Arg668Cys], p.[Ser977Phe;(TG)12T5] or p.[Ser977Phe;(TG)13T5], p.[Ser1235Arg;(TG)13T5]). Most patients (152/182, 83.52%) carried two mutations, 16 patients (8.79%) carried only one, whereas 14 patients (7.69%) were found with no *CFTR* alteration. Thirteen mutations were found in more than 1% of patients (Table 3), the most frequent being p.Phe508del (23.90%), IVS8-T5 (17.03%), p.[Asp443Tyr;Gly576Ala;Arg668Cys] (3.3%), p.Asp1152His (3.3%), p.Arg117His (3.02%) and p.Leu997Phe (3.02%). Fifty-eight mutations (15.9% of alleles) were found in only one patient. The most represented genotype is p.Phe508del *in trans* to IVS8-(TG)12T5 (30/182=16.5%). Two patients were found with p.Phe508del *in trans* to the complex p.[Ser1235Arg;(TG)13T5] allele. One sample harbored three mutations (p.Ser977Phe, IVS8-(TG)12T5 and IVS8-(TG)13T5); as no familial segregation was possible, we could not determine which IVS8-(TG)mTn allele was *in cis* to p.Ser977Phe. One CBAVD patient carried T3 at IVS8-Tn locus. Two patients originating from Maghreb were apparently homozygous, one for IVS8-(TG)11T5, the other one for the complex allele p.[Arg74Trp;Val201Met;Asp1270Asn]. Samples apparently homozygous for a mutation (these two cases) and samples with no or only one mutation (30 cases) were further screened for large rearrangements by SQF-PCR. Two large deletions were identified in two patients carrying another defect *in trans*: IVS8(TG)12T5 or p.Arg117His [17]. Overall, a *CFTR* defect was identified in 320/364 (87.9%) of *CFTR* alleles.

Molecular haplotyping of the IVS8-(TG)mTn polymorphic locus

inserm-00185153, version 1 - 10 Nov 2009

To validate our methodology for the determination of IVS8-(TG)mTn haplotypes, the fluorescent assay was applied to 50 CBAVD patients for which the Tn and (TG)m repeats had been previously determined by sequencing of exon 9/flanking regions. All the results were concordant, so that the remaining 132 CBAVD patients were analyzed using the fluorescent assay (Figure 2); alleles ranged from (TG)9 to (TG)13 repeats, associated with T3, T5, T7, or T9 alleles. The obtained patterns showed intra and inter-run reproducibility. When DNAs were available, we confirmed IVS8-(TG)mTn haplotypes by studying the familial segregation of alleles. Three (TG)m backgrounds were associated with IVS8-T5 alleles in this population of CBAVD samples: (TG)11-T5 (7/62, 11.3%), (TG)12-T5 (49/62, 79%) and (TG)13-T5 (4/62, 6.5%). For two patients the (TG)m repeats could not be determined due to DNA shortage. Of a total of 364 CBAVD alleles, 7 (1.92%) were IVS8-(TG)11T5: three patients carried this haplotype *in trans* to a CF mutation (p.Asn1303Lys, p.Val562Ile, or p.Arg1162X), one was homozygous and two patients carried only this haplotype. A rare IVS8-(TG)12T3 combination that we previously found *in trans* to p.Phe508Cys in one patient was confirmed [15].

SiCTA/sequencing assay validation

DNA from 10 patients already investigated for mutations by using our previous classic protocols for *CFTR* genetic testing were anonymized and submitted for analysis along with DNA from three wild-type subjects (CBAVD patient's partner negative for *CFTR* mutations). The panel included compound heterozygous samples, homozygous wild-type or mutant samples, and heterozygous samples with only one mutation. All the previously characterized alleles were identified, and no disease-causing mutations were detected in the normal controls. Apart from the validation phase, DNA from 58 additional patients (CBAVD, CF and suspected CF) were submitted to the analysis together with synthetical blood samples included in the European Molecular Quality Control schemes [external quality assessment

2005 scheme harboring 7 mutations (394delTT, p.Arg117His, p.Arg347His, IVS8-(TG)12T5, p.Ile507del, p.Arg553X, 2183AA>G); external quality assessment 2006 scheme harboring 9 mutations (394delTT, p.Arg117His, IVS8-(TG)12T5, p.Ile507del, p.Phe508Cys, p.Arg553X, 3876delA, 3905insT and p.Trp1282X)]. All the mutations were correctly detected. Twenty-eight patients previously analyzed by DGGE and/or DHPLC and for which only one or no mutation had been found were reanalyzed by SiCTA/sequencing: all the mutations and polymorphisms already detected were identified and no additional sequence change was found except polymorphisms in intronic regions that were not analyzed by the classic protocols.

inserm-00185153, version 1 - 10 Nov 2009

inserm-00185153, version 1 - 10 Nov 2009

Discussion

Three recent studies have reported extensive *CFTR* sequencing in 96-well plates in patients with classical or atypical cystic fibrosis. In two assays [24, 25] the *CFTR* gene was studied in 32 amplicons and each PCR primer contained a M13 linker sequence ensuring a single PCR condition and the use of universal priming in cycle sequencing. All PCR primers had to be redesigned due to the presence of the M13 linker sequence. In another assay [26], the *CFTR* gene was amplified in 30 amplicons with external primers and then sequenced using internal primers in 96-well plates. Redesigning all PCR primers was necessary and three amplification conditions had to be used due to different annealing temperatures of the primers.

The SiCTA/sequencing methodology described in this study presents several advantages: i) the technique is easy to set up as a routine, ii) it relies on the use of primers that have been used by diagnostic laboratories since the original description of the *CFTR* gene [18], which avoids risks of allele drop out due to the presence of single nucleotide polymorphism (SNP) in newly designed primers, iii) a touchdown PCR protocol enables single amplification conditions for all the separate amplicons, making the 96-well plate format possible and iv) the same primers are used both for PCR and sequencing. Moreover, the methodology is flexible as the *CFTR* genes of one to three CBAVD patients (32 to 96 amplicons) can be amplified simultaneously. The sequencing reactions can then be performed on a single plate either in one direction for three patients, or in both directions for one patient. Due to the 96-well plate format, both PCR and sequence reactions set-up can be performed with a pipetting robot or multichannel pipettors. All the polymorphisms detected by the classic protocols were found along with some others owing to the position of the primers; our assay extended a minimum of 50 to 200 bp into each intron at all intron-exon boundaries. Extensive sequencing has the advantage to characterize the polymorphisms in contrast with DGGE or DHPLC techniques that require an additional sequencing step of abnormal profiles including those induced by

inserm-00185153, version 1 - 10 Nov 2009

polymorphisms. The single condition touchdown amplification/sequencing strategy is now widely used in our laboratory e.g. for the comprehensive genotyping of 8 genes in the Usher syndrome for which up to 250 amplicons are analysed [27, 28].

The new fluorescent, rapid and reliable method that we report here for the determination of alleles and haplotypes at locus IVS8-(TG)mTn will also facilitate the genotyping of the sequences that modulate the splicing efficiency of exon 9 in the mRNA and/or are major determinants of the penetrance of the 5T allele in CBAVD. The importance of IVS8-(TG)mT5 determination is supported by an international collaborative study providing evidence that the odds of pathogenicity are 28 and 34 times greater for (TG)12T5 and (TG)13T5, respectively, than for (TG)11T5 [16]. The use of allele-specific oligonucleotide (ASO) hybridization, reverse hybridization or PCR with specific primers for the T5, T7 or T9 can result in misdiagnosis, not only in the assignation of the most common alleles (7, 9, or 5T) but also because of the inability to detect other variants than the most frequent T5, T7 and T9. Despite the fact that IVS8-T3 [10], T6 [11], T10 [12], T11 [13] and IVS8-(TG)8 [23] appear to be rare in Caucasian populations, our technique allows an easy detection of all different alleles described so far at IVS8-Tn (3, 5, 6, 7, 9, 10 and 11 thymidines) and -(TG)m (8 to 13 repeats) loci. Our fluorescent haplotyping method presents similar advantages as those previously described using a direct sequencing method [23]. It allows i) to determine IVS8-(TG)mTn haplotypes even without family linkage study, ii) to be used either as primary or confirmatory test. However, for familial segregation or population studies, the fluorescent haplotyping method is more rapid to implement and easier to interpret, thereby saving a considerable amount of time and effort. Direct molecular haplotyping of the IVS8-(TG)mTn repeats by melting curve analysis of hybridization probes was recently described [12]. This method is particularly rapid but, unfortunately, (TG)12T5 and (TG)13T5 haplotypes, that are

inserm-00185153, version 1 - 10 Nov 2009

of particular interest in CBAVD patients, showed indistinguishable melting temperatures and could not be clearly identified [29].

Large rearrangements account for 16-20% of unidentified alleles in classic cystic fibrosis [1, 4, 5] (our unpublished results). In our CBAVD series, only two patients (1.1%) have been found to carry a large gene deletion [17], which represents 4.8% of alleles negative for a *CFTR* point mutation, which is in accordance with the results of another study [7]. Gross alterations seem to be less frequent in CBAVD than in CF, which is relevant with the lower proportion of severe alleles in CBAVD than in CF. So far only eight cases with a large rearrangement have been reported in CBAVD [2, 5, 7, 17].

By combining direct sequencing, fluorescent haplotyping and SQF-PCR for large rearrangements, we were able to identify at least one CBAVD-causing mutation in 92.3% of the 182 patients analyzed in this series, including 83.5% with two alleles fully characterized, which is the highest rate reported so far. For the 16 patients with only one mutation, the implication of the *CFTR* gene is still unclear: do they carry a deep mutation in introns not investigated by techniques applied to genomic DNA? Is heterozygosity for a *CFTR* defect a predisposing factor for CBAVD? For the 14 patients negative for point mutation, large rearrangement, or predisposing haplotype, the link between CBAVD and *CFTR* gene is questionable.

The procedure can also be applied to partners of affected patients (CF or CBAVD) for whom the *CFTR* gene needs to be entirely scanned. Moreover, it is of particular interest for fetuses when hyperechogenic bowel or ascities suggestive of CF is prenatally detected by ultrasound during the 2nd or 3rd trimester of pregnancy. In such emergency situation, the common known mutations are first screened using a commercially available kit. If one of the parents carries a mutation that has been transmitted to the fetus, the entire *CFTR* gene has to be scanned for the

second mutation. Using SiCTA/sequencing, the whole gene can be analyzed for both known and private mutations within a few days.

inserm-00185153, version 1 - 10 Nov 2009

inserm-00185153, version 1 - 10 Nov 2009

Acknowledgements

We thank the University-Hospital of Montpellier and VLM (*Vaincre la Mucoviscidose*) for their support.

inserm-00185153, version 1 - 10 Nov 2009

References

1. Audrezet MP, Chen JM, Raguenes O, Chuzhanova N, Giteau K, Le Marechal C, Quere I, Cooper DN, Ferec C: Genomic rearrangements in the CFTR gene: extensive allelic heterogeneity and diverse mutational mechanisms. Hum Mutat 2004, 23:343-357
2. Niel F, Martin J, Dastot-Le Moal F, Costes B, Boissier B, Delattre V, Goossens M, Girodon E: Rapid detection of CFTR gene rearrangements impacts on genetic counselling in cystic fibrosis. J Med Genet 2004, 41:e118
3. Bombieri C, Bonizzato A, Castellani C, Assael BM, Pignatti PF: Frequency of large CFTR gene rearrangements in Italian CF patients. Eur J Hum Genet 2005, 13:687-689
4. Hantash FM, Redman JB, Starn K, Anderson B, Buller A, McGinniss MJ, Quan F, Peng M, Sun W, Strom CM: Novel and recurrent rearrangements in the CFTR gene: clinical and laboratory implications for cystic fibrosis screening. Hum Genet 2006, 119:126-136
5. Hantash FM, Milunsky A, Wang Z, Anderson B, Sun W, Anguiano A, Strom CM: A large deletion in the CFTR gene in CBAVD. Genet Med 2006, 8:93-95
6. Niel F, Legendre M, Bienvenu T, Bieth E, Lalau G, Sermet I, Bondeux D, Boukari R, Derelle J, Levy P, Ruszniewski P, Martin J, Costa C, Goossens M, Girodon E: A new large CFTR rearrangement illustrates the importance of searching for complex alleles. Hum Mutat 2006, 27:716-717
7. Ratbi I, Legendre M, Niel F, Martin J, Soufir JC, Izard V, Costes B, Costa C, Goossens M, Girodon E: Detection of cystic fibrosis transmembrane conductance regulator (CFTR) gene rearrangements enriches the mutation spectrum in congenital bilateral absence of the vas deferens and impacts on genetic counselling. Hum Reprod 2007

inserm-00185153, version 1 - 10 Nov 2009

8. Claustres M, Guittard C, Bozon D, Chevalier F, Verlingue C, Ferec C, Girodon E, Cazeneuve C, Bienvenu T, Lalau G, Dumur V, Feldmann D, Bieth E, Blayau M, Clavel C, Creveaux I, Malinge MC, Monnier N, Malzac P, Mittre H, Chomel JC, Bonnefont JP, Iron A, Chery M, Georges MD: Spectrum of CFTR mutations in cystic fibrosis and in congenital absence of the vas deferens in France. Hum Mutat 2000, 16:143-156
9. Claustres M: Molecular pathology of the CFTR locus in male infertility. Reprod Biomed Online 2005, 10:14-41
10. Buratti E, Baralle FE: Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. J Biol Chem 2001, 276:36337-36343
11. Viel M, Leroy C, Des Georges M, Claustres M, Bienvenu T: Novel length variant of the polypyrimidine tract within the splice acceptor site in intron 8 of the CFTR gene: consequences for genetic testing using standard assays. Eur J Hum Genet 2005, 13:136-138
12. Millson A, Pont-Kingdon G, Page S, Lyon E: Direct molecular haplotyping of the IVS-8 poly(TG) and polyT repeat tracts in the cystic fibrosis gene by melting curve analysis of hybridization probes. Clin Chem 2005, 51:1619-1623
13. Kobler D, Modi H, Goldman B: Identification of an 11T allele in the polypyrimidine tract of intron 8 of the CFTR gene. Genet Med 2006, 8:125-128
14. Cuppens H, Lin W, Jaspers M, Costes B, Teng H, Vankeerberghen A, Jorissen M, Droogmans G, Reynaert I, Goossens M, Nilius B, Cassiman JJ: 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 Invest 1998, 101:487-496

inserm-00185153, version 1 - 10 Nov 2009

15. Disset A, Michot C, Harris A, Buratti E, Claustres M, Tuffery-Giraud S: A T3 allele in the CFTR gene exacerbates exon 9 skipping in vas deferens and epididymal cell lines and is associated with Congenital Bilateral Absence of Vas Deferens (CBAVD). Hum Mutat 2005, 25:72-81

16. Groman JD, Hefferon TW, Casals T, Bassas L, Estivill X, Des Georges M, Guittard C, Koudova M, Fallin MD, Nemeth K, Fekete G, Kadasi L, Friedman K, Schwarz M, Bombieri C, Pignatti PF, Kanavakis E, Tzetis M, Schwartz M, Novelli G, D'Apice MR, Sobczynska-Tomaszewska A, Bal J, Stuhrmann M, Macek M, Jr., Claustres M, Cutting GR: Variation in a repeat sequence determines whether a common variant of the cystic fibrosis transmembrane conductance regulator gene is pathogenic or benign. Am J Hum Genet 2004, 74:176-179

17. Taulan M, Girardet A, Guittard C, Altieri JP, Templin C, Beroud C, des Georges M, Claustres M: Large genomic rearrangements in the CFTR gene contribute to CBAVD. BMC Med Genet 2007, 8:22

18. Zielenski J, Rozmahel R, Bozon D, Kerem B, Grzelczak Z, Riordan JR, Rommens J, Tsui LC: Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Genomics 1991, 10:214-228

19. Fanen P, Ghanem N, Vidaud M, Besmond C, Martin J, Costes B, Plassa F, Goossens M: Molecular characterization of cystic fibrosis: 16 novel mutations identified by analysis of the whole cystic fibrosis conductance transmembrane regulator (CFTR) coding regions and splice site junctions. Genomics 1992, 13:770-776

20. Chillon M, Dork T, Casals T, Gimenez J, Fonknechten N, Will K, Ramos D, Nunes V, Estivill X: 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

inserm-00185153, version 1 - 10 Nov 2009

fibrosis chromosomes and association with severe phenotype. Am J Hum Genet 1995, 56:623-629

21. Highsmith WE, Burch LH, Zhou Z, Olsen JC, Boat TE, Spock A, Gorvoy JD, Quittel L, Friedman KJ, Silverman LM, et al.: A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations. N Engl J Med 1994, 331:974-980

22. Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS: 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 1991, 19:4008

23. 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, 32:540-542, 544-547

24. Strom CM, Huang D, Chen C, Buller A, Peng M, Quan F, Redman J, Sun W: Extensive sequencing of the cystic fibrosis transmembrane regulator gene: assay validation and unexpected benefits of developing a comprehensive test. Genet Med 2003, 5:9-14

25. 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

26. McGinniss MJ, Chen C, Redman JB, Buller A, Quan F, Peng M, Giusti R, Hantash FM, Huang D, Sun W, Strom CM: Extensive sequencing of the CFTR gene: lessons learned from the first 157 patient samples. Hum Genet 2005, 118:331-338

27. Baux D, Larrieu L, Blanchet C, Hamel C, Ben Salah S, Vielle A, Gilbert-Dussardier B, Holder M, Calvas P, Philip N, Edery P, Bonneau D, Claustres M, Malcolm S, Roux

inserm-00185153, version 1 - 10 Nov 2009

AF: Molecular and in silico analyses of the full-length isoform of usherin identify new pathogenic alleles in Usher type II patients. Hum Mutat 2007

28. Roux AF, Faugere V, Le Guedard S, Pallares-Ruiz N, Vielle A, Chambert S, Marlin S, Hamel C, Gilbert B, Malcolm S, Claustres M: Survey of the frequency of USH1 gene mutations in a cohort of Usher patients shows the importance of cadherin 23 and protocadherin 15 genes and establishes a detection rate of above 90%. J Med Genet 2006, 43:763-768

29. Costa C, Goossens M, Girodon E: Simultaneous molecular haplotyping of both IVS8 (TG)m and (T)n tracts in the CFTR gene: still a challenge. Clin Chem 2006, 52:1621-1622; author reply 1622

30. Chillon M, Casals T, Mercier B, Bassas L, Lissens W, Silber S, Romey MC, 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

31. Dork T, Dworniczak B, Aulehla-Scholz C, Wieczorek D, Bohm I, Mayerova A, Seydewitz HH, Nieschlag E, Meschede D, Horst J, Pander HJ, Sperling H, Ratjen F, Passarge E, Schmidtke J, Stuhrmann M: Distinct spectrum of CFTR gene mutations in congenital absence of vas deferens. Hum Genet 1997, 100:365-377

inserm-00185153, version 1 - 10 Nov 2009

Legend for figures

Figure 1: Strategy and position of the primers for IVS8-(TG)m and -Tn repeats analysis and (TG)mTn haplotyping.

A – Position and sequence of the primers used for exon 9 and flanking regions amplification (first PCR).

B – Focus on IVS8-(TG)mTn polymorphic region: position and sequence of the primers used for the 3 internal fluorescent PCRs. Upper case: 5' end sequence of exon 9; lower case: 3' end sequence of IVS8 containing the (TG)m and Tn repeats; in bold: primers used to determine the IVS8-(TG)mTn haplotypes; star: fluorescent label.

Figure 2: IVS8-(TG)mTn analysis

A – Electropherograms of IVS8-(TG)m, -Tn repeats and (TG)mTn haplotypes for two patients ([(TG)9T9]+[(TG)12T5] and [(TG)10T7]+[(TG)11T7]). The peaks of interest are shown in black.

B – Example of five different IVS8-(TG)m repeats. Due to the sequence context IVS8-(TG)m repeats are slippage prone. From our experience, the interpretation rule is to consider only the highest and the last peak (in black).

inserm-00185153, version 1 - 10 Nov 2009

Table 1: Conditions for PCR amplification of *CFTR* exon 9, IVS8-(TG)m or -Tn repeats and IVS8-(TG)mTn haplotypes.

	Primer name	**Sequence 5'->3'**	**Reference**	**Amplicon length**	**Annealing temperature (°C) / number of cycles**
Exon 9	9i5	TAATGGATCATGGGCCATGT	[18]	558 bp for (TG)11T7	56 / 20
	9i3	ACAGTGTTGAATGTGGTGCA	[18]		
(TG)m	CF9	TGAAAATATCTGACAAACTC	[19]	67 bp for (TG)11	55 / 22
	RI8TG8	(6FAM)GCGGCGGAAACACACACACACACACACA	This study		
Tn	I9D3T	CCGCCGCTGTGTGTGTGTGTGTGTGTTT	Adapted from [30]	118 bp for T7	55 / 22
	IVS8Rm	(HEX)CTGAAGAAGAGGCTGTCATC	This study		

Underlined: stabilizating tail

(TG)mTn (CF9-IVS8Rm) = 152 bp for (TG)11-T7; 55°C/28 cycles

inserm-00185153, version 1 - 10 Nov 2009

Table 2: Primers used for PCR amplification and sequencing of the *CFTR* gene.

Location	Forward sequence (5'-> 3')	Reference	Reverse sequence (5'-> 3')	Reference
Ex 1	CGTAGTGGGTGGAGAAAGC	[19]	AGCGCATCTTTTTAAAAACTGCTTA	This study
Ex 2	CAAATCTGTATGGAGACC (a)	This study	TGTTTGCTTTCTCTTCTCTAAAT	This study
Ex 3	CTTGGGTTAATCTCCTTGGA	[18]	ATTCACCAGATTTCGTAGTC	[18]
Ex 4	TCACATATGGTATGACCCTC	[18]	TTGTACCAGCTCACTACCTA	[18]
Ex 5	ATTTCTGCCTAGATGCTGGG	[18]	AACTCCGCCTTTCCAGTTGT	[18]
Ex 6a	TTAGTGTGCTCAGAACCACG	[18]	CTATGCATAGAGCAGTCCTG	[18]
Ex 6b	TGGAATGAGTCTGTACAGCG (b)	[18]	GAGGTGGAAGTCTACCATGA	[18]
	GATTTACAGAGATCAGAG (b)	This study		
Ex 7	AGACCATGCTCAGATCTTCCAT	[18]	GCAAAGTTCATTAGAACTGATC	[18]
Ex 8	TGAATCCTAGTGCTTGGCAA	[18]	TCGCCATTAGGATGAAATCC	[18]
Ex 9	TAATGGATCATGGGCCATGT (b)	[18]	ACAGTGTTGAATGTGGTGCA	[18]
	TTTTTAACAGGGATTTGGGG (b)	This study		
Ex 10	GCAGAGTACCTGAAACAGGA	[18]	CATTCACAGTAGCTTACCCA	[18]
Ex 11	CAACTGTGGTTAAAGCAATAGTGT	[18]	GCACAGATTCTGAGTAACCATAAT	[18]
IVS11	TTTCTTAATTGTGTGCTGAATAC (a)	This study	CAGTTCCCATATTAAATAGAAATGA	[20]
Ex 12	GTGAATCGATGTGGTGACCA	[18]	CTGGTTTAGCATGAGGCGGT	[18]
Ex 13(5') (c)	TGCTAAAATACGAGACATATTGCA	[18]	ATCTGGTACTAAGGACAG	[18]
Ex 13(3') (c)	TCAATCCAATCAACTCTATACGAA	[18]	TACTCCTTATCCTAATCCTATGAT	[18]
Ex 14a	AAAAGGTATGCCACTGTTAAG	[18]	GTATACATCCCCAAACTATCT	[18]
Ex 14b	GAACACCTAGTACAGCTGCT	[18]	TACATACAAACATAGTGGATT	This study
Ex 15	TCAGTAAGTAACTTTGGCTGC	[19]	CCTATTGATGGTGGATCAGC (d)	[19]
Ex 16	CAGAGAAATTGGTCGTTACT	[18]	ATCTAAATGTGGGATTGCCT	[18]
Ex 17a	TGCAATGTGAAAATGTTTAC	This study	TGTACACCAACTGTGGTAAG	[18]
Ex 17b	TTCAAAGAATGGCACCAGTGT	[18]	ATAACCTATAGAATGCAGCA	[18]
Ex 18	GTAGATGCTGTGATGAACTG	[18]	AGTGGCTATCTATGAGAAGG	[18]
Ex 19	GCCCGACAAATAACCAAGTGA	[18]	GCTAACACATTGCTTCAGGCT	[18]

IVS19	GAATCATTCAGTGGGTATAACCAG	[21]	AGGCTTCTCAGTGATCTCTTG	[21]
Ex 20	GGTCAGGATTGAAAGTGTGCA	[18]	CTATGAGAAAACTGCACTGGA	[18]
Ex 21	AATGTTCACAAGGGACTCCA	[18]	CAAAAGTACCTGTTGCTCCA	[18]
Ex 22	AAACGCTGAGCCTCACAAGA	[18]	TGTCACCATGAAGCAGGCAT	[18]
Ex 23	AGAAGTACTGGTGATTCTAC	[31]	TAAAGCTGGATGGCTGTATG	[18]
Ex 24	GGACACAGCAGTTAAATGTG	[18]	ACTATTGCCAGGAAGCCATT	[18]

(a) The forward primers of exon 2 and IVS11 were adapted from Zielenski *et al.* [18] and Chillon *et al.* [20] respectively.

(b) The 5’ end of exons 6b and 9 were sequenced using 2 different forward primers to avoid the polymorphic regions (IVS6a-(GATT)n and IVS8-(TG)mTn tracts respectively).

(c) Exon 13 was amplified in 2 overlapping amplicons Ex 13(5’) and Ex13(3’) due to its large size.

(d) The reverse primer of exon 15 designed by Fanen *et al.* [19] was deprived of the GC tail.

inserm-00185153, version 1 - 10 Nov 2009

inserm-00185153, version 1 - 10 Nov 2009

Table 3: Mutations of the *CFTR* gene represented in more than 1% of CBAVD patients.

Mutation	No. of alleles	% of the 364 alleles
p.Phe508del [a,b,c]	87	23.90
IVS8-T5 [a,b]	62	17.03
p.[Asp443Tyr;Gly576Ala;Arg668Cys]	12	3.30
p.Asp1152His	12	3.30
p.Arg117His [a,b,c]	11	3.02
p.Leu997Phe	11	3.02
p.Gly542X [a,b,c]	9	2.47
p.Leu206Trp	7	1.92
p.Met952Ile	6	1.65
p.Arg347His [a]	5	1.37
p.[Arg74Trp;Val201Met;Asp1270Asn]	5	1.37
p.Arg170His	4	1.10
p.Phe508Cys	4	1.10

Mutations that are part of commercial kit panels: [a] Cystic Fibrosis v3 5/7/9T OLA ASR, Abbott diagnostics, USA; [b] INNO-LiPA *CFTR*, Innogenetics, Belgium and [c] Elucigene CF30, Tepnel diagnostics, UK.