Human Molecular Genetics, 2003, Vol. 12, No. 10 **1111–1120**
DOI: 10.1093/hmg/ddg131

New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12

Franco Pagani[1], Cristiana Stuani[1], Maria Tzetis[2], Emmanuel Kanavakis[2], Alexandra Efthymiadou[2], Stavros Doudounakis[3], Teresa Casals[4] and Francisco E. Baralle[1,*]

[1]International Centre for Genetic Engineering and Biotechnology, Padriciano 99, Trieste 34012, Italy, [2]Department of Medical Genetics, Athens University, 'Aghia Sophia' Children's Hospital, Thivon and Livadias, Athens 11527, Greece, [3]Cystic Fibrosis Department, 'Aghia Sophia' Children's Hospital, Athens, Greece and [4]Centro de Genética Medica y Molecular, Instituto Recerca Oncológica (IRO), Hospital Duran y Reynals, Gran Via s/n km 2,7 08907 Barcelona, Spain

Received December 16, 2002; Revised and Accepted March 17, 2003

The increase in genome scanning data, derived from clinical genetics practice, is producing a wealth of information on human sequence variability. The critical issue is to identify if a given nucleotide change results in a benign polymorphism or a disease-causing mutation. We have focused on one specific gene expression step, pre-mRNA processing, where we can functionally define the effect of nucleotide changes and in turn the patient's mutation can shed light on the basic pre mRNA splicing mechanisms. Our results show that several nucleotide changes in CFTR exon 12 induce a variable extent of exon skipping that leads to reduced levels of normal transcripts. This is the case in both natural mutations D565G and G576A (the latter having previously considered a neutral polymorphism) and several site-directed silent substitutions. We demonstrate here that this phenomenon is due to the interference with a new regulatory element that we have named composite exonic regulatory element of splicing (CERES). The effect of single nucleotide substitutions at CERES cannot be predicted by neither SR matrices nor enhancer identification. The recognition and characterization of splicing abnormalities, caused by exon sequence variations at CERES elements, may represent a frequent disease-causing mechanism that also relates to the phenotypic variability. Our results indicate that even the most benign looking polymorphism in an exon cannot be ignored as it may affect the splicing process. Hence, appropriate functional splicing assays should be included in genotype screenings to distinguish between polymorphisms and pathogenic mutations.

INTRODUCTION

Pre-mRNA splicing is a complex mechanism that relies on the correct identification on the transcribed RNA of protein coding sequences (exons) from the more abundant non coding sequences (introns). This identification requires the presence of a 'core' of *cis*-acting sequences (5′ and 3′ splice sites, branch-point sequence and polypyrimidine tract) together with additional *cis*-acting elements, localized either in the coding sequence or in nearby introns, that may enhance or antagonize splicing. All these elements, interacting with splicing regulatory factors such as the serine-arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) (1–4), allow the fine tuning of the splicing process and provide the possibility of its differential regulation. A special case of modulation is the alternative splicing, a strictly developmentally regulated and tissue-specific controlled phenomenon that occurs in 40–60% of human genes and is responsible for significant genome diversity (5,6). However, such a complexity entails a high risk of derangement following even minor mutations. Abnormalities in pre-mRNA splicing actually represent an important mechanism by which gene mutations cause disease, accounting for 30–50% of cases when RNA was directly analysed (7,8). In classical splicing defects the core elements

*To whom correspondence should be addressed. Tel: +39 0403757337; Fax: +39 0403757361; Email: baralle@icgeb.org

described above are mutated and the effect on splicing can be predicted on the basis of the genomic DNA sequence analysis. However, the effect of mutations in non-obvious regulatory elements is more difficult to evaluate directly from genome scanning methodologies. Furthermore, they may result in less severe splicing defects and less severe phenotypes. This will require the determination of the RNA expression pattern in affected tissues that may not always be available. In particular, point mutations in exonic regulatory elements may induce variable levels of aberrant exon skipping and thus the distinction between polymorphisms and splicing pathogenic mutations represents an increased challenge (4,9,10). To facilitate this distinction the use of computer assisted enhancer sequence prediction by means of SR protein score matrices has been suggested (9,11). In addition, the frequency by which exonic sequence variations may affect the splicing process is unknown.

To understand the role of non-canonical regulatory elements in splicing pathology we have analysed the splicing pattern resulting from variants at the CFTR exon 12. Skipping of this exon removes a highly conserved region encoding part of the first nucleotide-binding fold of CFTR rendering the protein non-functional (12). Complete skipping, due to disruption at the conserved splice sites, causes severe classical cystic fibrosis (CF) with typical multiple organ involvement (13–15). In normal individuals, variable levels of CFTR transcripts without exon 12, accounting for 5–30% of total CFTR mRNA, have been observed (16–18). We have focused our attention on the effect on splicing of several previously reported changes in coding sequences, that did not have an obvious association with protein functionality or disease phenotype. This is the case for two interesting and enigmatic missense mutations, D565G and G576A. The D565G mutation was previously reported in a young subject during a screening program and suspected of inducing exon skipping (19). The conservative G576A missense substitution was originally listed as a neutral polymorphism in the Cystic Fibrosis Genetic Analysis Consortium (http://genet.sickkids.on.ca). Later studies detected the mutation in individuals with classical CF (20) and in patients who have evidence of a clinical disease only in a subgroup of the organ systems (20–22). These non-classical CF forms, including late-onset pulmonary disease, congenital bilateral absence of vas deferens (CBAVD), or idiopathic pancreatitis frequently show a genetic diagnostic challenge for the unclear genotype–phenotype correlation (22–24). In this paper we have analysed *in vivo* and *in vitro* the splicing defects associated with exon 12 nucleotide variations. For this purpose we studied mRNA transcripts derived from patients' nasal epithelial cells and from hybrid minigene transient transfection assays. Our results demonstrate that a similar splicing pattern is observed for a given mutation both *in vivo* and *in vitro*. Several nucleotide substitutions in CFTR exon 12 change the splicing pattern and define two new splicing elements with a composite regulatory function that we have named CERES (composite exon regulatory element of splicing). Mutations inducing exon skipping include both missense D565G and G576A mutants and several neutral substitutions. The splicing defect of neutral substitutions suggests that careful interpretation is needed for the pathological nature of nucleotide variations in human disease, particularly when they do not change the amino acid sequence.

RESULTS

D565G and G576A missense mutations cause CFTR exon 12 skipping *in vivo*

We evaluated, in nasal epithelial cells, the pattern of CFTR exon 12 splicing in both normal subjects and heterozygous individuals with D565G and G576A alleles. The missense D565G mutation was detected in seven Greek subjects, always in *cis* with the common polymorphism R668C (2134C/T) in exon 13 (Table 1). All these subjects were heterozygotes and two showed non-classical CF: one patient was affected by CBAVD and the other with pulmonary symptoms of an unknown aetiology. The patient carrying the G576A missense mutation was affected by testicular azoospermia and in this case the G576A allele was also in *cis* with the R668C polymorphism. To distinguish between the transcripts produced from the normal and mutant alleles we took advantage of the presence of the R668C polymorphism in exon 13 in *cis* with both mutations, and designed allele-specific primers. The 688C and 688R primers contain either C or T at their 3′ end (nucleotide 2134 in CFTR mRNA), to discriminate between the R and C alleles at amino acid 668 (Fig. 1A). Two PCR's were set up for the nasal epithelial cell cDNA derived from each of the D565G and G576A heterozygotes and from heterozygous controls for R668C, using the common F3 forward primer in exon 11 and each of the two allele specific primers of exon 13 (Fig. 1A). Analysis of the cDNA products in all cases revealed the presence of two transcripts of 449 and 362 bp containing or lacking the exon 12, respectively (Fig. 1B). In heterozygous individuals the 668C allele carrying the mutations D565G or G576A clearly showed a significantly lower proportion of normal transcripts containing exon 12 than the 668R allele (Fig. 1B, lanes 7–20). On the contrary, in normal subjects, the two polymorphic alleles with 668C or 668R produced an equally low amount of leaky splicing (Fig. 1B, lanes 1–6). The presence of leaky splicing in normal alleles is consistent with previous data where, in normal individuals, variable amounts of mRNA transcripts lacking exon 12 (5–30%) have been reported (16). In order to quantitate the proportion of exon 12+ CFTR mRNA accurately, *in vitro* transcribed mRNAs, with and without exon 12, were mixed in varying proportions, reverse transcribed and analysed as for the nasal samples (Fig. 1C, left panel). The data were then plotted against the proportion of input RNAs (Fig. 1C, right panel) and, according to the resulting graph, the experimental proportions of CFTR exon 12 transcripts in nasal epithelial cells were corrected. This analysis showed that the mutant D565G and G576A alleles produced about 40 and 22% of exon inclusion, respectively (Table 1). These results indicate that, in nasal epithelial cells, the D565G and G576A missense mutations cause a splicing defect affecting the recognition of CFTR exon 12.

Defective CFTR exon 12 recognition in hybrid minigenes containing D565G and G576A missense mutations

In order to study in more detail the splicing regulation of CFTR exon 12 we have developed a faithful splicing assay that mimics the *in vivo* situation. Exon 12 sequences including part of the

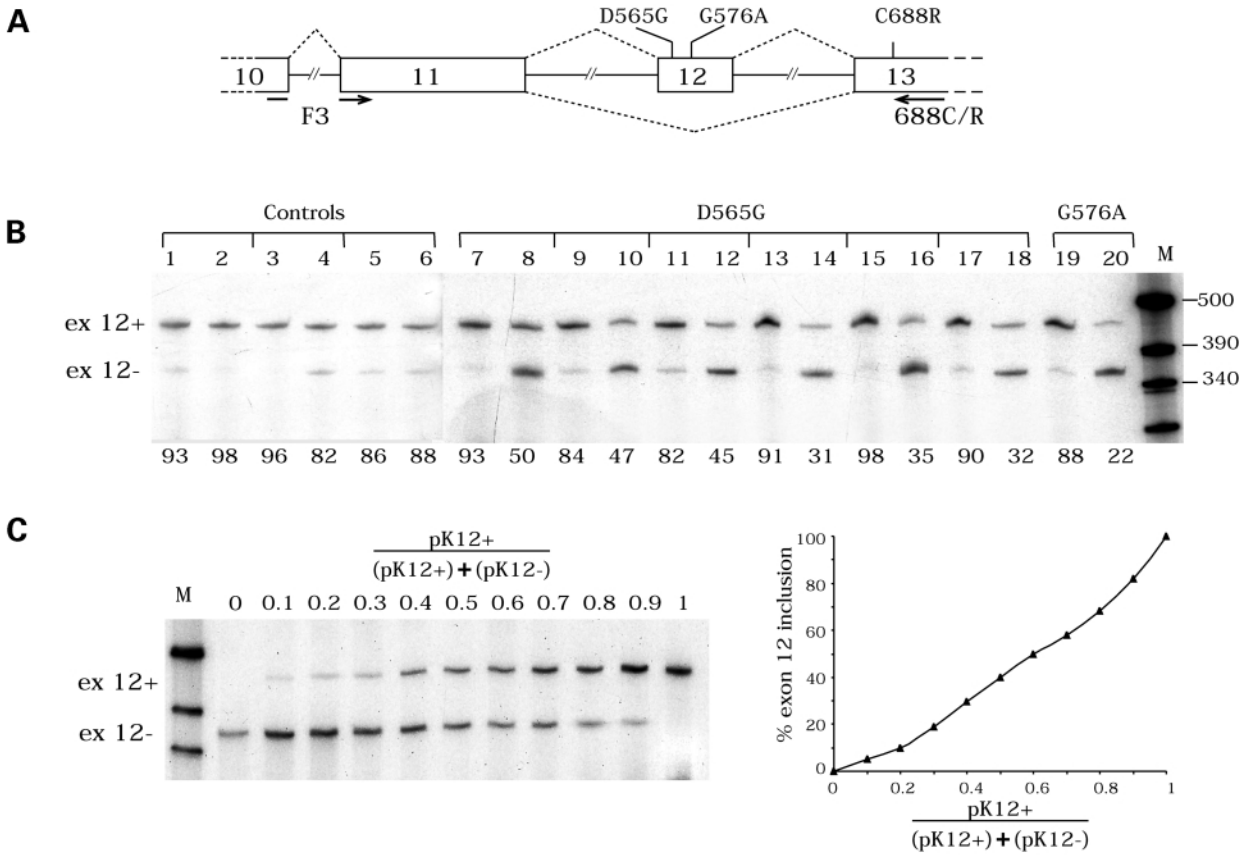

Figure 1. RT–PCR allele-specific amplification experiments in D565G and G576A carriers. (**A**) Genomic organization of the CFTR gene. The position of the missense substitutions (D565G and G576A) in exon 12 and of the C668R polymorphic variant in exon 13 is indicated. CFTR exons are boxed, introns are represented as solid lines and the splicing patterns as dotted lines. The oligonucleotides used in allele specific PCR are indicated as arrows. The D565G and G576A carriers presented the missense mutation in *cis* with the 668C variant. (**B**) RT–PCR allele-specific products analysed in 5% denaturing acrylamide gel. RNA extracted from nasal epithelial cells from R688C heterozygous controls (lanes 1–6), from D565G carriers (lanes 7–18) and from the G576A carrier (lanes 19–20), was reverse transcribed and amplified with F3/668C primers (even-numbered lanes) and with F3/668R primers (odd-numbered lanes). F3/668C primers detect the processed transcrips originating from mutant allele in patients and the normal allele in the controls, respectively. The F3/668R primers detect the processed transcripts originating from the normal allele in both patients and controls. The transcripts of 449 and 362 bp with and without the exon 12 are shown. M, molecular weight marker. The specificity of the 668R and 668C primers was evaluated by amplification experiments in homozygous R668C individuals. In this experimental conditions the primers did not amplify non-specific alleles (data not shown). Numbers below each lane represent the percentage of exon 12 inclusion corrected according to the graph shown in (**C**) quantitative RT–PCR analysis of CFTR exon 12. Left panel: 5% denaturing acrylamide gel of mixture templates. *In vitro* transcribed RNAs from pK12+ and pK12− (with the arginine variant) were mixed in different proportions, reverse transcribed and amplified with primers F3 and 688R. Left panel: the CFTR exon 12 plus and minus products analysed in 5% PAGE were quantitated using a Cyclone Instant Imager. The percentage of inclusion calculated as the ratio of the intensity of the upper and lower bands (normalized for the GC content) was plotted against the ratio of pK12+ RNA to pK12− RNA plus pK12+ RNA. Data represent the mean of three experiments done in duplicate. Similar percentage of exon 12 inclusion was obtained using pK12 RNAs with the Cysteine variant, amplified with the allele specific 688C primer.

Table 1. Allele-specific PCR transcript analysis

	F3/668R	F3/668C
Subjects with D565G mutation	89.7 ± 5.9	40 ± 8.3[a]
Subject with G576A mutation	88 ± 3	22 ± 4[b]
Normal controls (heterozygotes for polymorphism R668C)	91.7 ± 5.1	89.3 ± 8.1

Data from six subjects with the D565G mutation, one patient with the G576A mutation and four controls are calculated from the experimental proportions of CFTR exon 12 inclusion adjusted according to the graph shown in Figure 1C. Data are expressed as percentage of exon 12 inclusion and represent the average ± SD of at least two independent measurements done in duplicate.
[a]Transcripts containing exon 12 derived from the D565G allele.
[b]Transcripts containing exon 12 derived from the G576A allele.

flanking introns 11 and 12 were inserted into a previously reported hybrid minigene consisting of the α-globin and fibronectin intron-exon sequences (25) (Fig. 2A). This minigene has been widely used to investigate the regulation of normal and aberrant splicing of several exons including CFTR exon 9 (25) and the cryptic exon in the *ATM* gene (26). Transfection of normal WT CFTR exon 12 minigene in Hep3B cells followed by RT–PCR amplification generated mostly transcripts of 326 bp containing exon 12 and a smaller proportion of 239 bp transcripts in which it was lacking (Fig. 2). The same was true for WT-B, a variant with a new restriction site in intron 12 created to facilitate subsequent cloning procedures (Fig. 2C, lanes 1–2). Quantitative analysis of the results showed that about 17% of the mRNA did not contain exon 12, indicating a leaky splicing of CFTR exon 12 minigene analogous to that of the

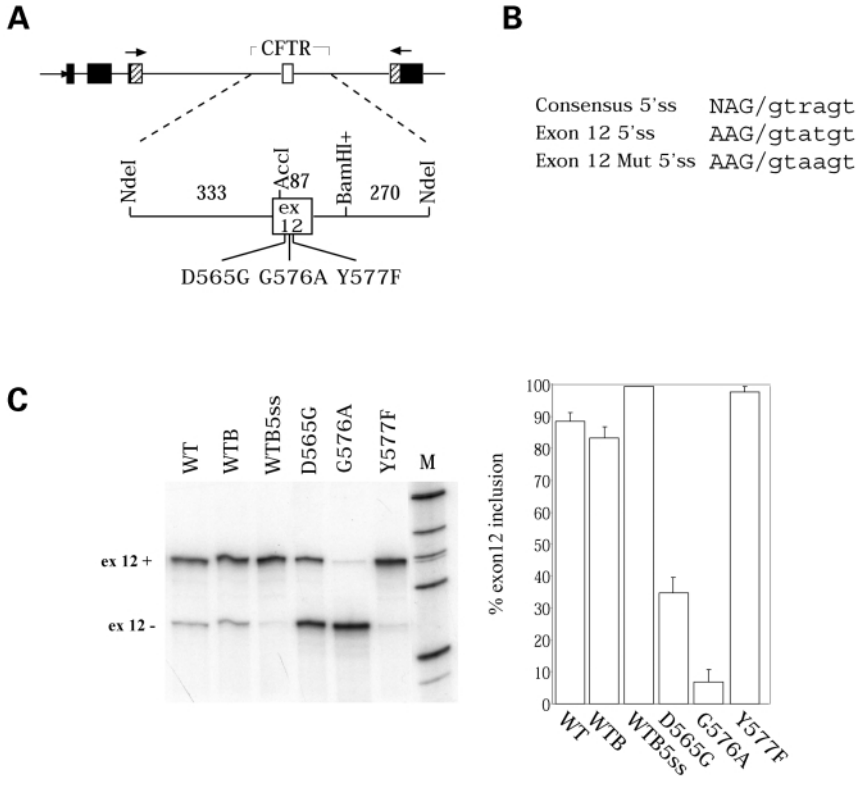

Figure 2. CFTR exon 12 hybrid minigene transient transfection assay. (**A**) Schematic representation of the hybrid CFTR exon 12 minigenes used in transient transfection splicing assay. α-Globin, fibronectin EDB and human CFTR exons are indicated in black, shaded and white boxes, respectively. The primers used in RT–PCR assay are indicated as superimposed arrows. The length of the CFTR fragments (intron 11, exon 12 and intron 13) along with the position of relevant restriction sites (including the *BamH*I restriction site created to facilitate subsequent cloning procedures) and of the missense mutations in the exon are indicated. (**B**) Sequence comparison of wild type exon 12 5′ splice site with the consensus. The underlined T to A substitution in Mut 5′ss restores the consensus sequence in the WTB5ss minigene. (**C**) pre-mRNA processing of the human CFTR exon 12 minigenes generated transcripts with inclusion (exon 12+) and exclusion (exon 12−) of the exon. The indicated minigene variants were transfected in Hep3B cells and the splicing pattern analyzed by RT–PCR amplification. The left panel is a radioactive 6% acrylamide gel electrophoresis of the RT–PCR products. Quantitation of the percentage of CFTR exon 12 inclusion (right panel) was performed with a Cyclone Instant Imager, and the results corrected for the CG content of the splicing variants. Each lane represents the mean ± SD of at least three independent experiments done in duplicate.

chromosomal gene in tissues from normal individuals. The reason for the leaky splicing may be a deviation from the consensus 5′ splice site at a critical T in position +4 (Fig. 2B). In fact, the leaky splicing of the exon can be corrected by restoring the consensus adenine in this position which results in complete exon inclusion (Fig. 2C, lane WTB5ss).

We then studied the pattern of splicing of a minigene with the missense mutations D565G, G576A and Y577F, the latter associated to classical CF. The D565G showed about 35% of normal transcripts containing exon 12 (Fig. 2C). The G576A mutation resulted in a severe splicing defect, with only 7% of normal exon 12+ mRNA transcripts. Instead, the nearby Y577F mutation found in classical CF did not produce aberrant skipping but surprisingly increased exon inclusion in comparison with WTB (Fig. 2C). This indicates that, unlike G565A and D565G, the disease-causing effect of Y577F cannot be attributed to a splicing abnormality but rather to a protein defect. These results indicate that the two missense D565G and G576A mutations associated with non-classical CF induce variable proportion of exon 12 skipping, with G576A most severely affected. The disease-causing effect of this last conservative mutation must hence be due to the induced splicing defect.

Negative regulation of CFTR exon 12 splicing by the splicing factors SF2/ASF and hnRNPA1

In several gene systems tissue-specific factors may modulate the splicing efficiency (1,3,25). To evaluate their role in CFTR exon 12 we transfected normal and the three D565G, G576A and Y577F minigenes in different cell lines (Fig. 3A). For each cell line tested, the three variants cause comparable changes in splicing efficiency, with D565G and G576A inducing exon skipping and Y577F exon inclusion. However, the severity of the splicing defect was different among the cell lines. For example, G576A showed 15% of exon inclusion in NT2, 7–9% in HeLa, COS1 and T84, and complete skipping in CFPAC. These results suggest that the intracellular concentration of ubiquitous splicing factors may modulate the severity of the splicing defects. To test this hypothesis we have evaluated the role of several splicing factors using normal and mutated CFTR exon 12 minigenes. Figure 3B shows the results of the cotransfection experiments with two representative splicing factors: the SF2/ASF that belongs to the family of the SR proteins and the heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1). Plasmids coding for these splicing factors were

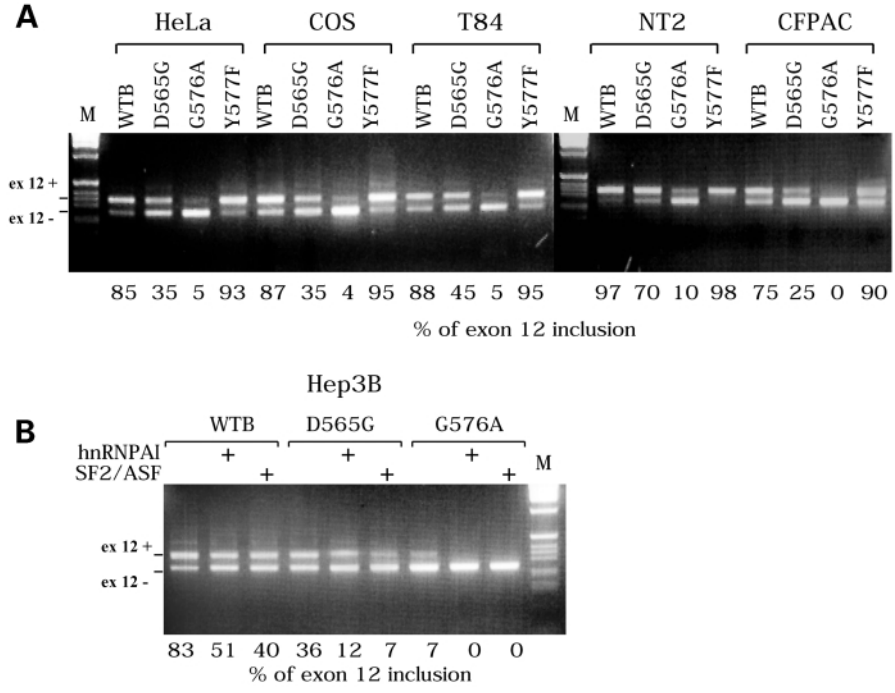

Figure 3. Expression of the CFTR exon 12 missense variants in different cell lines and in the presence of regulatory splicing factors. (**A**) CFTR exon 12 hybrid minigenes were transfected in HeLa, COS, T84, NT2 and CFPAC cell lines and the amplified products loaded on 1.5% agarose gel. (**B**) Minigene variants were transfected in Hep3B cells along with 500 ng of the empty control pCG vector or the indicated splicing factor plasmids. RNA splicing variants were detected by RT–PCR and analysed on a 1.5% agarose gels. Exon 12 positive (+) and negative (−) mRNAs are indicated. The number below each lane indicates the percentage of exon 12 inclusion and represents the mean of two independent experiments done in duplicate.

transiently expressed simultaneously with the CFTR constructs and their effect on splicing analysed by RT–PCR amplification. In WTB, D565G and G576A, overexpression of any of the two splicing factors caused an increase in CFTR exon 12 skipping. The amount of normal transcript containing the exon 12 were reduced in WTB (40–51%), very low in D565G (7–12%) and virtually absent in the G576A mutant (Fig. 3B). A similar inhibition of splicing was observed for SRp55, SRp40 and SRp75 splicing factors, whereas polypyrimidine tract binding protein did not change the splicing pattern (data not shown). These results indicate that, in the presence of high concentration of inhibitory splicing factors, the D565G and G576A missense mutations produce very low levels of normal CFTR transcripts. Variations in the concentration of these regulatory splicing factors may be responsible for tissue-specific phenotypic expression.

Definition of composite exonic regulatory elements of splicing

To understand the frequency with which exonic point variations affect the splicing process we have prepared several variants with single nucleotide changes at or close to those of the missense mutations. A total number of 26 hybrid minigenes were analysed containing site-directed mutations at two target sequences of the exon: the AAGATGC sequence at the 5′ end from position 12 to 18, which includes D565G at position 15, and a central GGATAC sequence from position 47 to 52 which contains G576A and Y577F of position 48 and 51, respectively (Fig. 4). The variant minigenes were transfected into Hep3B cells and the resulting splicing patterns analysed. Unexpectedly, all but one of the 26 mutations affected the splicing process. Thirteen caused severe exon skipping (less than 15% of exon inclusion), five affected the splicing process mildly (between 60 and 15% of exon inclusion) and seven enhanced the exon inclusion (Fig. 5).

In the AAGATGC element all mutations at positions 14, 15, 17 and 18 induced variable degrees of exon skipping (Fig. 5). Base replacements at position 16 increased exon 12 inclusion and substitutions at position 12 showed diametrically opposite effects on splicing depending on the base changed: 12G induced inclusion (98%) and 12T skipping (8%; Fig. 5). Similarly, substitutions at position 13 resulted in either increased (13G) or reduced (13T) exon recognition. A similar picture was evident at the central regulatory sequence GGATAC in which site-directed mutants may cause different degrees of exon skipping (positions 47, 48, 49 and 52), exon inclusion (position 51) or variable changes depending on the base substituted (position 50) (Fig. 5). In some exons, SR protein score motifs have been found to correlate with splicing efficiency, therefore raising the possibility of predicting the exon skipping effect of single base substitutions (9,11). We did not find any significant correlation with exon 12 skipping using the four available motif-scoring matrices (data not shown). This indicates that the two exonic sequences analysed contain composite elements with overlapping enhancer and silencer properties that are not completely dependent on the SR protein interactions. Owing to the peculiar behaviour, we named these

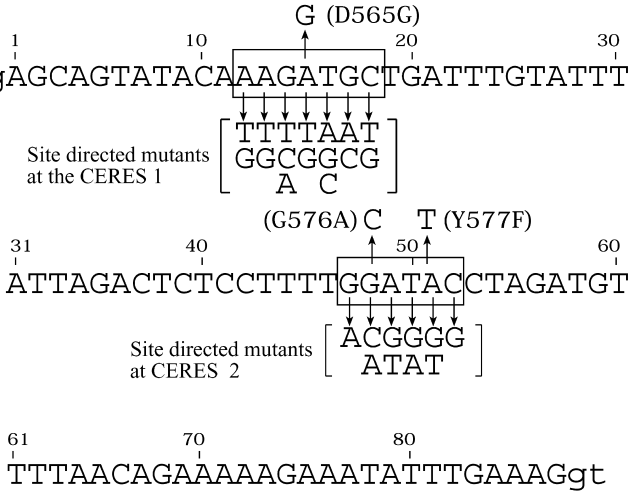

Figure 4. Nucleotide sequence of the CFTR exon 12 (in uppercase) showing the position of the two CERES (boxed) and the site-directed mutants. Donor and acceptor splice sites are in lower case. The four site-directed mutants that do not change the aminoacid code are underlined. The missense mutations in the one-letter code are indicated in brackets.

elements composite exonic regulatory element of splicing (CERES).

Interestingly, the site-directed mutants 13G, 16C, 49G and 49T are at the third position of codons and do not modify the aminoacid code. So they would be labelled as neutral variations, if found in a classical genome scanning analysis. Among these, the base changes at position 49 resulted in very low amounts of normal transcripts containing exon 12 (about 10%; Fig. 5), indicating that 'neutral' variations frequently considered not to have pathological consequences may on the contrary induce a severe splicing defect.

DISCUSSION

Genotype screening in human disease frequently identifies exon sequence variations whose association with the disease phenotype is unclear. In fact the pathologic effect of an apparently benign polymorphism, such as codon third position variations or conservative amino acid replacement, are difficult to assess. In this paper we have evaluated the role of the exon coding sequence variability on the pre-mRNA splicing process, using the CFTR exon 12 as a model. We show that the disease-causing mechanism of several exonic substitutions involve the splicing process, resulting in a variable extent of exon skipping. The largely unexpected high frequency of the nucleotide changes that affects CFTR exon 12 splicing, (half of the changes resulted in less than 15% of exon inclusion, Fig. 5), has wider implications for all genotype work and for the pre mRNA splicing field. These results indicate the possibility that even the most benign looking polymorphism in an exon can modify the clinical phenotype by causing aberrant splicing. An attempt to provide a tool for the prediction of splicing defects using SR consensus sequences has been successful in some systems (9,11). However this procedure did not show significant correlation in CFTR exon 12 (this paper) or in CFTR exon 9 (unpublished).

In order to identify 'hidden' splicing mutations and to characterize the largely unknown exonic regulatory elements, we have analysed transcripts derived from nasal epithelial cells of CFTR patients and extended this analysis with hybrid minigene assays. (4,25–27). The mild CF phenotypes of the D565G and G576A patients may be explained considering that the variant CFTR protein is functional and that the defect is the consequence of exon 12 skipping induced by the nucleotide change. It is interesting to note that changes in the codon following G576A result in the classical CF mutation Y577F that produces a non-functional protein, but splicing is not significantly altered. Two similar splicing defects that produce a variable amount of normal CFTR transcripts (skipping of exon 9 or inclusion of a cryptic exon by 3849 + 10 Kb C→T), are indeed associated with tissue-specific and phenotype variability (23,24,28,29). These variants produce different quantities of normal mRNAs in different tissues from the same individual (30,31), suggesting a tissue-specific modulation by splicing factors. In the same manner, it should not be considered that the missense mutations in CFTR exon 12 produce constant proportions of normal mRNAs *in vivo*. In fact, their splicing efficiency can be modulated by the cell type (Fig. 3A), possibly through variable concentrations of the splicing factors such as SF2/ASF and hnRNPA1 (Fig. 3B). Thus, according to tissue concentration of regulatory splicing factors, and their variations from individual to individual, the D565G and G576A may produce different quantities of mRNAs lacking the exon leading to phenotype variability. The identification of additional regulatory *trans*-acting factors involved in splicing modulation and the determination of their relative concentrations in affected tissues will give a more comprehensive picture of the molecular interactions that directly influence the phenotypic expression.

Exonic regulatory splicing elements found in both constitutive and alternative exons are classically divided in exonic splicing enhancers and silencers (4,9,10). In *CFTR* exon 12, the systematic site directed mutagenesis experiments clearly show overlapping enhancer and silencer functions and therefore the CERES element is neither a pure enhancer nor a silencer (Fig. 5). Highly degenerated exonic enhancer elements have been suggested to be fully predictable using scoring matrices that allow the direct testing of point mutations on splicing, and therefore may have a practical application in human pathology (9,11). In the CFTR exon 12 CERES, the exon skipping phenotype was not predicted by the available computer-assisted analysis of SR sites. It should be noted that work on enhancers and silencers has mostly involved limited mutagenesis and that extensive nucleotide changes around these elements may uncover a CERES type behaviour in many currently accepted pure 'enhancers' or 'silencers'. The differences between classical and CERES-type regulatory elements may not only be due to the lack of extensive mutagenesis studies but more importantly to the fact that the classical enhancers and silencers are mostly defined using the simplified *in vitro* splicing systems. These systems do not necessarily reflect the nuclear cell environment where transcription and RNA processing are tightly coordinated (32–34). An additional property that should differentiate classical enhancers–silencers and CERES is the capacity of the former to promote autonomously splicing changes in heterologous context, a property that we did not observe for the CERES (data not shown). The CERES

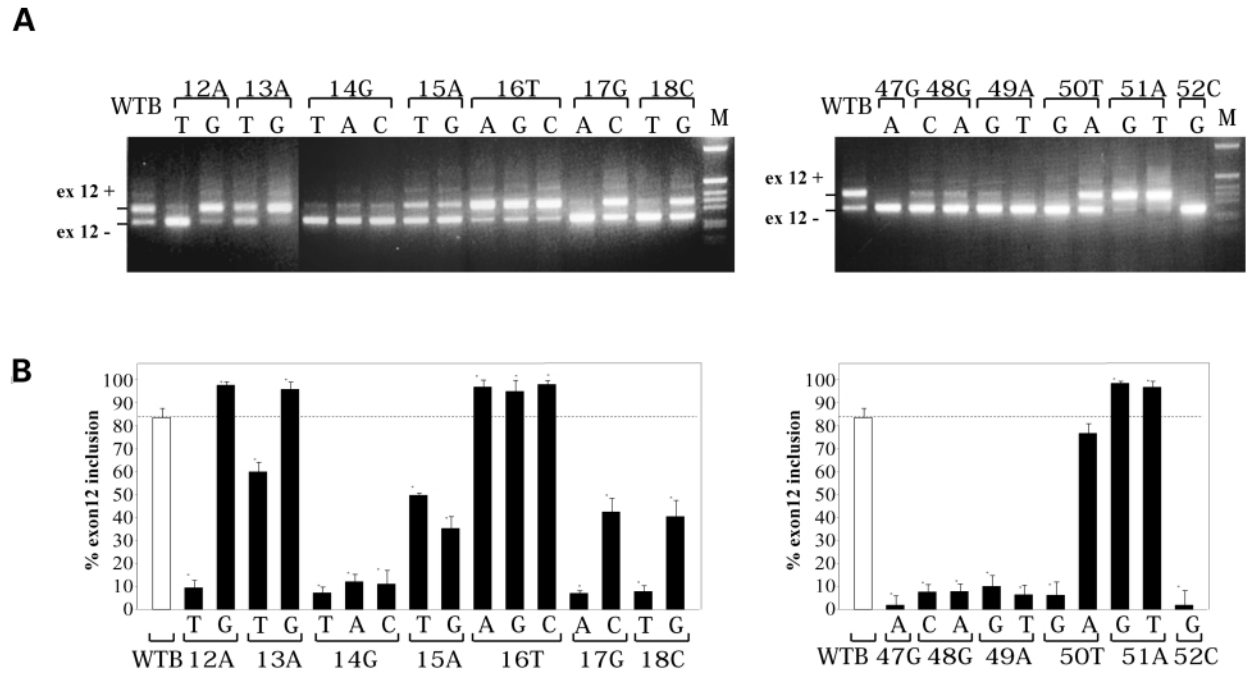

Figure 5. Effect of the site-directed mutants at the two composite exonic regulatory elements of splicing in CFTR exon 12. (**A**) Agarose gel electrophoresis of RT–PCR products from splicing assay from the site-directed mutants in the AAGATGC (left panel) and GGATAC (right panel) CERES. Hep3B cells were transfected with 3 μg of the indicated variants. (**B**) The histogram shows the percentage of exon 12+ and exon 12− transcripts of the site-directed mutants detected by radioactive PCR and quantitated using a Cyclone Instant Imager. Each bar represents the mean ± SD of three independent experiments done in duplicate.

elements of the CFTR exon 12 seem to be strongly dependent on the context for its function and, accordingly, the composite characteristics of these elements may also extend to the flanking nucleotides.

The identification of CERES with similar overlapping enhancer and silencer functions in other exons, such as the CFTR exon 9 (unpublished data), indicates a common splicing regulatory role of these elements. Their function may be more evident in those exons with a weak 5′ splice site, such as the CFTR exon 12. The presence of multiple clusters of CERES elements distributed over the entire exon, or in relatively discrete regions, may produce an exon-specific combination of trans-acting factors targets and/or RNA structures that correctly define the entire sequence of CFTR exon 12. These types of mechanisms have been separately been identified in the determination of the efficiency of exon recognition in other systems (4,9,11,35,36).

The sequence composition of the splicing regulatory elements in CFTR exon 12 overlap with the codon usage preferences and requirements for protein function. Some of the site-directed mutants that induce a high amount of aberrant exon skipping are at the third position of the codon usage and do not change the aminoacid code (Fig. 4). When these variants are found in genomic screening, their effect on splicing is very rarely assessed and their location in splicing regulatory elements may be overlooked. Similarly the effect of missense mutations on exon recognition may also be unnoticed. An analysis of mRNA splicing pattern when possible *in vivo* and/or by employing reliable minigene *in vivo* splicing assays, should be mandatory both for proper genetic diagnosis and for improvement of our knowledge on basic RNA processing mechanisms.

MATERIALS AND METHODS

Patients and DNA mutation analysis

Nasal epithelial cells were collected from six individuals carriers of mutation D565G (A>G at 1826 in CFTR cDNA), from one CBAVD patient with G576A and from four non-CF control individuals heterozygotes for the polymorphism R668C. None of the controls used in this study had clinical signs of typical CF. Clinical data and CFTR genotypes of all the D565G and G576A carriers are shown in Table 2. Genomic DNA was extracted from white blood cells, as previously described (37). The presence of mutations in the 27 exons and neighbouring intronic regions of the CFTR gene was assessed by denaturing gradient gel electrophoresis (DGGE) following simple or multiplex PCR of patient DNA (38–40). All DNA samples showing a shift in DGGE mobility and not presenting a pattern of a known mutation were sequenced using an automatic DNA sequencer (Vistra, model 725-Molecular Dynamics) for the identification of the mutation. The phase of linkage for missense mutations D565G and G556A and polymorphism R668C was deduced from family studies and confirmed by sequencing of allele specific cDNAs.

Allele-specific amplification analysis and quantitative PCR

Nasal epithelial cells were collected after brushing the inferior turbinates using interdental brushes (ParoIsola, Thalwil, Switzerland), immediately put into extraction buffer and RNA prepared by the RNeasy method (Qiagen, Hilden,

Table 2. Data on individual carriers of D565G and G576A mutations

Patient code	CFTR genotype[a]	Age (years)	Sex	Reason for CF testing	Sweat test (mEq/l)	Pancreatic status
1PM	D565G/−	16	F	Nasal polyposis, Sa	65, 70	PS
2DF	D565G − 1717−9T>C/−	7	F	Recurrent episodes of pneumonia	41.4	PS
3MA	D565G/−	Adult	M	Carrier status	nt	nt
4KA	D565G/−	Adult	F	Carrier status	nt	nt
5KP	D565G/−	Adult	M	Carrier status	nt	nt
6PRA	D565G/−	Adult	F	Carrier status	nt	nt
7ORA[b]	D565G/−	Adult	M	CBAVD	<40	PS
8	G576A/−	Adult	M	Testicular azoospermia	nt	nt

PS, pancreatic sufficiency; *Sa*, *Staphylococcus*; nt, non-tested; CBAVD, congenital bilateral absence of vas deferens. All patients are heterozygous for the R668C allele.
[a]A minus sign denotes absence of CFTR mutation after DGGE analysis of all 27 exons and adjoining intronic sequences; N denotes normal allele.
[b]RNA sample not available.

Germany) according to the manufacturer's instructions. First-strand cDNA was synthesized using hexanucleotide primers and MULV reverse transcriptase (GeneAmp Kit, Perkin Elmer). After the first-strand cDNA synthesis two PCRs were set up both with the same forward F3 5′-gaagaggacatctccaagtttgca-gag-3′ primer and either 668C 5′-ggagcatctccttctaatgagaaacg-3′ or 668R 5′-ggagcatctccttctaatgagaaaca-3′ reverse primers, amplifying either the 668C or the 668R alleles, respectively. Reverse transcription–polymerase chain reaction (RT–PCR) was performed using Hot Star polymerase (Quiagen) according to the manufacturer's instructions, which includes an initial denaturation step at 95°C for 15 min for the activation of the polymerase. Thirty cycles PCR (95°C for 30 s, 66°C for 45 s, and extension at 72°C for 1 min) were used for the amplification of CFTR cDNA. [α^{32}]dCTP was included in the PCR reaction mixture, the products loaded on 5% denaturing polyacrylamide–8 M urea gel, dried and exposed to a Cyclone Instant Imager. Each analysis was repeated at least twice on two independent RNA preparations. The specificity of the allele amplification assay was verified by PCR amplification of cDNA samples derived from homozygotes for the variants R or C in position 668.

For quantitative PCR analysis, cDNA amplified fragments from nasal epithelia with and without exon 12 and with the 668R or 668C variants, were blunt cloned in the *Sma*I site of the pBSKS vector under the T7 Promoter to obtain pK12+ Arg, pK12− Arg, pK12+ Cys and pK12− Cys. Transcription of cold RNA was done as previously described (41). T7 transcribed RNA from pK12+ Arg (or pK12+ Cys) was mixed in varying proportions with pK12− Arg (or pK12− Cys) T7 transcribed RNAs. These template mixtures were reverse-transcribed and amplified by 30 cycle PCRs with primers F3 and 688R (or 688C) in the presence of [α^{32}]dCTP and the resulting products were analysed in 5% PAGE and quantitated by a Cyclone Instant Imager. The percentage of inclusion calculated as the ratio of the intensity of the upper and lower bands (normalized for the GC content) was plotted against the ratio of pK12+ RNA to pK12− RNA plus pK12+ RNA. A similar percentage of exon inclusion was obtained for each point for either pK12 Arg or pK12 Cys amplified RNAs with 688R or 688C primers, respectively.

Hybrid minigene constructs

Wild-type CFTR sequences consisting of the last 333 bases of intron 11, exon 12 (87 bp) and the first 270 bases of intron 12 were amplified from normal genomic DNA with exon 12 dir 5′-gccatatgctatggtacagttcagt-3′ and exon 12 rev 5′-catatgcta-catagatagcaattgctataac-3′ primers. These primers contained *Nde*I site at their ends that were used to clone the CFTR exon 12 variants in the unique *Nde*I site of pTBNde(min). pTBNde(min) represents the basic hybrid minigene construct utilised and consists of a modified version of the α-globin-fibronectin EDB minigene (25). In pTBNde(min) transcription is driven from the minimal α-globin promoter and SV40 enhancer (25). By PCR mediated site-directed mutagenesis with ex12 Bam Rev 5′-aaatggatcctatgatgggacagtc-3′ and exon 12 Bam Dir 5′atcataggatccattttacctcttgag-3′ we introduced a *BamH*I restriction site 71 bp downstream of the 5′ splice site in intron 12 to facilitate subsequent cloning procedures creating the WTB minigene. The missense and artificial point mutations were introduced in this minigene between the *Acc*I and *BamH*I sites. This region was substituted with the appropriate *Acc*I–*BamH*I cassettes created by PCR-mediated site directed mutagenesis. The oligonucleotides used for PCR-mediated mutagenesis are available upon request. The exon 12 and flanking intronic sequences of all hybrid minigenes were verified by sequence analysis.

Analysis of the hybrid minigene expression

Hep3B (human hepatocarcinoma), HeLa (human cervical carcinoma), COS 1 (SV40-transformed simian kidney), CFPAC-1 (ductal pancreatic adenocarcinoma derived for a CF patient homozygote for ΔF508) and NT2D1 (human embrinal teratocarcinoma) cells were grown in Dulbecco's modified Eagle's medium supplemented with 4.5 g/l glucose, 10% fetal calf serum, 50 μg/ml gentamicin and 4 nM glutamine. T84 (colon carcinoma) cells were maintained Dulbecco's modified Eagle's F-12HAM medium supplemented with 4.5 g/l glucose, 10% fetal calf serum, 50 μg/ml gentamicin and 4 nM glutamine. Plasmid DNA was purified with JetStar columns (Genomed, Wielandstrasse, Germany). Cells were transfected with the DOTAP reagents with 3 μg of each

reported plasmid and the control empty vector pCG (0.5 μg) or different amounts of the two splicing factors codifying plasmids for SF2/ASF and hnRNPA1. RNA extraction was performed after 48 h according to the method of Chomczynski and Sacchi (42). cDNA was synthesized with hexanucleotide random primers using MoMuLV reverse transcriptase (Gibco BRL, Grand Island, NY, USA). RT–PCR was done as previously described (25) with the primers 2-3α 5′-caacttcaagctcctaagccactgc-3′ and B2 5′-taggatccggtcaccaggaagttggttaaatca-3′ (35 cycles, 45 s at 94°C, 45 s at 58°C, 45 s at 72°C), using 2U of *Taq* DNA polymerase (Roche). PCRs were optimised to remain in the exponential range of amplification and products were routinely fractionated in 1.5% (w/v) agarose gel. For quantitation of the PCR reactions, [α^{32}]dCTP was included in the PCR reaction mixture, the products loaded on 5% denaturing polyacrylamide–8 M urea gel, dried and exposed to a Cyclone Instant Imager. The counts of each splicing band were corrected by the number of C/G present in the PCR-product sequence.

ACKNOWLEDGEMENTS

We thank A. Krainer for the SR protein score motifs, T. Doerk and R. Garcia for helpful comments and discussion, Sara Larriba for technical assistance and Ann Crum for proof-reading the manuscript. The financial support of Telethon-Italy (grant no. GGP02453) is gratefully acknowledged.

REFERENCES

1. Mayeda, A. and Krainer, A.R. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*, **68**, 365–375.
2. Zahler, A.M., Neugebauer, K.M., Lane, W.S. and Roth, M.B. (1993) Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science*, **260**, 219–222.
3. Caceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R. (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science*, **265**, 1706–1709.
4. Muro, A.F., Caputi, M., Pariyarath, R., Pagani, F., Buratti, E. and Baralle, F.E. (1999) Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. *Mol. Cell Biol.*, **19**, 2657–2671.
5. Consortium, I.H.G.S. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
6. Modrek, B. and Lee, C. (2002) A genomic view of alternative splicing. *Nat. Genet.*, **30**, 13–19.
7. Ars, E., Serra, E., Garcia, J., Kruyer, H., Gaona, A., Lazaro, C. and Estivill, X. (2000) Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum. Mol. Genet.*, **9**, 237–247.
8. Teraoka, S.N., Telatar, M., Becker-Catania, S., Liang, T., Onengut, S., Tolun, A., Chessa, L., Sanal, O., Bernatowska, E., Gatti, R.A. and Concannon, P. (1999) Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *Am. J. Hum. Genet.*, **64**, 1617–1631.
9. Liu, H.X., Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2001) A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat. Genet.*, **27**, 55–58.
10. Caputi, M., Casari, G., Guenzi, S., Tagliabue, R., Sidoli, A., Melo, C.A. and Baralle, F.E. (1994) A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucl. Acids Res.*, **22**, 1018–1022.
11. Cartegni, L. and Krainer, A.R. (2002) Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat. Genet.*, **30**, 377–384.
12. Delaney, S.J., Rich, D.P., Thomson, S.A., Hargrave, M.R., Lovelock, P.K., Welsh, M.J. and Wainwright, B.J. (1993) Cystic fibrosis transmembrane conductance regulator splice variants are not conserved and fail to produce chloride channels. *Nat. Genet.*, **4**, 426–431.
13. Strong, T.V., Smit, L.S., Nasr, S., Wood, D.L., Cole, J.L., Iannuzzi, M.C., Stern, R.C. and Collins, F.S. (1992) Characterization of an intron 12 splice donor mutation in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Hum. Mutat.*, **1**, 380–387.
14. Hull, J., Shackleton, S. and Harris, A. (1993) Abnormal mRNA splicing resulting from three different mutations in the CFTR gene. *Hum. Mol. Genet.*, **2**, 689–692.
15. Zielenski, J., Markiewicz, D., Lin, S.P., Huang, F.Y., Yang-Feng, T.L. and Tsui, L.C. (1995) Skipping of exon 12 as a consequence of a point mutation (1898 + 5G→T) in the cystic fibrosis transmembrane conductance regulator gene found in a consanguineous Chinese family. *Clin. Genet.*, **47**, 125–132.
16. Bremer, S., Hoof, T., Wilke, M., Busche, R., Scholte, B., Riordan, J.R., Maass, G. and Tummler, B. (1992) Quantitative expression patterns of multidrug-resistance P-glycoprotein (MDR1) and differentially spliced cystic-fibrosis transmembrane-conductance regulator mRNA transcripts in human epithelia. *Eur. J. Biochem.*, **206**, 137–149.
17. Slomski, R., Schloesser, M., Berg, L.P., Wagner, M., Kakkar, V.V., Cooper, D.N. and Reiss, J. (1992) Omission of exon 12 in cystic fibrosis transmembrane conductance regulator (CFTR) gene transcripts. *Hum. Genet.*, **89**, 615–619.
18. Hull, J., Shackleton, S. and Harris, A. (1994) Analysis of mutations and alternative splicing patterns in the CFTR gene using mRNA derived from nasal epithelial cells. *Hum. Mol. Genet.*, **3**, 1141–1146.
19. Tzetis, M., Efthymiadou, A., Doudounakis, S. and Kanavakis, E. (2001) Qualitative and quantitative analysis of mRNA associated with four putative splicing mutations (621 + 3A→G, 2751 + 2T→A, 296 + 1G→C, 1717-9T→C-D565G) and one nonsense mutation (E822X) in the CFTR gene. *Hum. Genet.*, **109**, 592–601.
20. Bienvenu, T., Adjiman, M., Thiounn, N., Jeanpierre, M., Hubert, D., Lepercoq, J., Francoual, C., Wolf, J., Izard, V., Jouannet, P., Kaplan, J.C. and Beldjord, C. (1997) Molecular diagnosis of congenital bilateral absence of the vas deferens: analyses of the CFTR gene in 64 French patients. *Ann. Genet.*, **40**, 5–9.
21. Ravnik-Glavac, M., Dean, M. and Glavac, D. (2000) Study of mutant and polyvariant mutant CFTR genes in patients with congenital absence of the vas deferens. *Pflugers Arch.*, **439**, R53–55.
22. Pignatti, P.F., Bombieri, C., Marigo, C., Benetazzo, M. and Luisetti, M. (1995) Increased incidence of cystic fibrosis gene mutations in adults with disseminated bronchiectasis. *Hum. Mol. Genet.*, **4**, 635–639.
23. Chillon, M., Casals, T., Mercier, B., Bassas, L., Lissens, W., Silber, S., Romey, M.C., Ruiz-Romero, J., Verlingue, C., Claustres, M., Nunes, V., Ferec, C. and Estivill, X. (1995) Mutations in the cystic fibrosis gene in patients with congenital absence of the vas deferens. *New Engl. J. Med.*, **332**, 1475–1480.
24. Costes, B., Girodon, E., Ghanem, N., Flori, E., Jardin, A., Soufir, J.C. and Goossens, M. (1995) Frequent occurrence of the CFTR intron 8 (TG)n 5T allele in men with congenital bilateral absence of the vas deferens. *Eur. J. Hum. Genet.*, **3**, 285–293.
25. Pagani, F., Buratti, E., Stuani, C., Romano, M., Zuccato, E., Niksic, M., Giglio, L., Faraguna, D. and Baralle, F.E. (2000) Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element. *J. Biol. Chem.*, **275**, 21041–21047.
26. Pagani, F., Buratti, E., Stuani, C., Bendix, R., Dork, T. and Baralle, F.E. (2002) A new type of mutation causes a splicing defect in ATM. *Nat. Genet.*, **30**, 426–429.
27. Niksic, M., Romano, M., Buratti, E., Pagani, F. and Baralle, F.E. (1999) Functional analysis of cis-acting elements regulating the alternative splicing of human CFTR exon 9. *Hum. Mol. Genet.*, **8**, 2339–2349.
28. Kerem, E., Rave-Harel, N., Augarten, A., Madgar, I., Nissim-Rafinia, M., Yahav, Y., Goshen, R., Bentur, L., Rivlin, J., Aviram, M. *et al.* (1997) A cystic fibrosis transmembrane conductance regulator splice variant with partial penetrance associated with variable cystic fibrosis presentations. *Am. J. Respir. Crit. Care Med.*, **155**, 1914–1920.
29. Chiba-Falek, O., Kerem, E., Shoshani, T., Aviram, M., Augarten, A., Bentur, L., Tal, A., Tullis, E., Rahat, A. and Kerem, B. (1998) The molecular basis of disease variability among cystic fibrosis patients carrying the 3849 + 10 kb C→T mutation. *Genomics*, **53**, 276–283.

30. Larriba, S., Bassas, L., Gimenez, J., Ramos, M.D., Segura, A., Nunes, V., Estivill, X. and Casals, T. (1998) Testicular CFTR splice variants in patients with congenital absence of the vas deferens. *Hum. Mol. Genet.*, **7**, 1739–1743.
31. Chiba-Falek, O., Parad, R.B., Kerem, E. and Kerem, B. (1999) Variable levels of normal RNA in different fetal organs carrying a cystic fibrosis transmembrane conductance regulator splicing mutation. *Am. J. Respir. Crit. Care Med.*, **159**, 1998–2002.
32. Bentley, D. (1999) Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell Biol.*, **11**, 347–351.
33. Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499–506.
34. Pagani, F., Stuani, C., Zuccato, E., Kornblihtt, A.R. and Baralle, F.E. (2002) Promoter architecture modulates CFTR Exon 9 skipping. *J. Biol. Chem.*, **5**, 5.
35. D'Souza, I. and Schellenberg, G.D. (2000) Determinants of 4-repeat tau expression. Coordination between enhancing and inhibitory splicing sequences for exon 10 inclusion. *J. Biol. Chem.*, **275**, 17700–17709.
36. Varani, L., Hasegawa, M., Spillantini, M.G., Smith, M.J., Murrell, J.R., Ghetti, B., Klug, A., Goedert, M. and Varani, G. (1999) Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc. Natl Acad. Sci. USA*, **96**, 8229–8234.
37. Miller, S.A., Dykes, D.D. and Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl. Acids Res.*, **16**, 1215.
38. Fanen, P., Ghanem, N., Vidaud, M., Besmond, C., Martin, J., Costes, B., Plassa, F. and Goossens, M. (1992) 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*, **13**, 770–776.
39. Costes, B., Fanen, P., Goossens, M. and Ghanem, N. (1993) A rapid, efficient, and sensitive assay for simultaneous detection of multiple cystic fibrosis mutations. *Hum. Mutat.*, **2**, 185–191.
40. Tzetis, M., Kanavakis, E., Antoniadi, T., Doudounakis, S., Adam, G. and Kattamis, C. (1997) Characterization of more than 85% of cystic fibrosis alleles in the Greek population, including five novel mutations. *Hum. Genet.*, **99**, 121–125.
41. Buratti, E., Dork, T., Zuccato, E., Pagani, F., Romano, M. and Baralle, F.E. (2001) Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J.*, **20**, 1774–1784.
42. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, **162**, 156–159.