*Published online 12 November 2009*

*Nucleic Acids Research, 2010, Vol. 38, No. 2* ***647–659***
*doi:10.1093/nar/gkp1040*

# Functional properties and evolutionary splicing constraints on a composite exonic regulatory element of splicing in CFTR exon 12

**Ariful Haque, Emanuele Buratti and Francisco E. Baralle***

International Centre for Genetic Engineering and Biotechnology (ICGEB), 34149 Trieste, Italy

Received September 8, 2009; Revised and Accepted October 22, 2009

## ABSTRACT

**In general, splicing regulatory elements are defined as Enhancers or Silencers depending on their positive or negative effect upon exon inclusion. Often, these sequences are usually present separate from each other in exonic/intronic sequences. The Composite Exonic Splicing Regulatory Elements (CERES) represent an extreme physical overlap of enhancer/silencer activity. As a result, when CERES elements are mutated the consequences on the splicing process are difficult to predict. Here, we show that the functional activity of the CERES2 sequence in CFTR exon 12 is regulated by the binding, in very close proximity to each other, of several SR and hnRNP proteins. Moreover, our results show that practically the entire exon 12 sequence context participate in its definition. The consequences of this situation can be observed at the evolutionary level by comparing changes in conservation of different splicing elements in different species. In conclusion, our study highlights how it is increasingly difficult to define many exonic sequences by simply breaking them down in isolated enhancer/silencer or even neutral elements. The real picture is close to one of continuous competition between positive and negative factors where affinity for the target sequences and other dynamic factors decide the inclusion or exclusion of the exon.**

## INTRODUCTION

Only a few years ago, mutations in the protein-coding section of genes that did not affect amino acid coding capacities were considered to be neutral with regards to the protein functional properties (and thus the evolutionary fitness of the gene). Since then, advances in both the pre-mRNA splicing and the translational research fields have severely challenged this assumption, especially with regards to its implications in the occurrence of human disease and on evolutionary mechanisms in general (1–3). With regards to the pre-mRNA alternative splicing process (4), this assumption was first challenged by the discovery that splicing regulatory elements (SREs) could be found embedded within exonic coding sequences both in alternative and constitutive splicing (5,6) and that these elements and the strength of the basic splicing consensus motifs (7) could regulate exon inclusion levels. More than 20 years since these discoveries, the importance of splicing regulatory regions within exon coding sequences has, if anything, greatly increased. In fact, global analyses of splicing events (8,9) and the search for these splicing regulatory motifs embedded within exons has received considerable attention, especially through the use of high-throughput and bioinformatics approaches (10,11). The results of all these analyses clearly indicate that SRE elements represent important players in controlling both alternatively and constitutively splicing processes (12,13).

In general, SRE elements are individually referred to as Exonic and Intronic Splicing Enhancers (ESEs and ISEs) or Exonic and Intronic Splicing Silencers (ESSs and ISSs) sequences depending on their localization and functional effects (14–18). The way in which SRE elements exert their action is through the binding of *trans*-acting factors, predominantly belonging to the SR and hnRNP protein families (19–23). It should be noted, however, that the list of proteins capable of modulating splicing is growing every year (24) and much still remains to be uncovered in this area of research. Nonetheless, aside from individual identities, it has been determined long ago that the balance between antagonistic factors binding to a particular SRE element is one key factor in discriminating exon inclusion/exclusion levels (25,26).

*To whom correspondence should be addressed. Tel: +39 040 3757337; Fax: +39 040 3757361; Email: baralle@icgeb.org

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

The possibility of controlling this balance has provided a great advantage to the eukaryotic cell because, whenever necessary, it can be shifted in one direction or the other through several mechanisms. In fact, beside the intrinsic affinity of splicing factors for their respective *cis*-targets, the binding level of each factor can be easily modified by variations in its relative expression levels (27–29), posttranslational modifications (30–33) or local RNA folding arrangements that limit/enhance their availability (34–39).

Studying these issues also makes for fascinating insights with regards to the potential relationships between coding and splicing regulatory regions during the course of evolution. It is now clear, in fact, that synonymous or even advantageous substitutions at the protein level may still be selected against if they end up to be harmful with regards to splicing decisions. On the other hand, suboptimal codon usage arrangements may be maintained to preserve correct splicing functionality (1,40). In keeping with this concept, recent analyses have uncovered the presence of extensive purifying selection against substitutions in ESE elements as determined by the reduced single nucleotide polymorphism density in these regions (41) and by the fact that codon usage at the exon–intron boundaries may be considerably affected by the need to maintain SRE sequences (42). It should also be mentioned, however, that these kind of analyses are still hampered by our limited knowledge of SRE sequences and caution should be employed when making these kind of comparisons on the basis of bioinformatic studies (43). For this reason, it is advisable to back up any eventual conclusion with functional experiments that might either support or not the bioinformatics considerations.

Among the known SRE sequences, a particularly interesting class of elements is represented by the discovery of the Composite Exonic Regulatory Elements of Splicing (CERES), that were first identified in human CFTR exon 12 and exon 9 (44,45). Unlike the classical exonic regulatory elements, that tend to predominantly possess either enhancing or silencing properties, the effect of mutagenesis in CERES elements is very unpredictable with regards to splicing outcomes. This makes it very difficult to evaluate the potential pathologic effect of apparently benign substitutions in these regions. In this work, we have combined the analysis of natural pathogenic mutations with a comparative human–mouse genomic approach to better understand the characteristics of one of these CERES elements in CFTR exon 12 (CERES2). The results of this analysis have reinforced the emerging concept that in many cases dividing exonic sequences in well defined enhancer/silencer or neutral splicing regulatory elements does not satisfactorily explain anymore the effects of artificial and natural substitutions. It is only by adopting a more global view of splicing regulatory codes that will allow us to understanding many dynamic sequence interplays aimed at preserving splicing definition of eukaryotic exons.

## MATERIALS AND METHODS

### Hybrid minigene constructs

Human CFTR exon 12 minigene constructs (T40C, G48C, A49G, A51T, C52T and WT) have been previously described (40,44). Further modification of the exon were introduced by PCR-directed mutagenesis using specific primers and cloned inside the NdeI restriction site of the pTB plasmid. Primer sequences for each described mutants can be provided up on request. Mouse CFTR Exon 12 along with the mouse introns (297 nt of intron 11 and 219 nt of intron 12) was amplified from genomic DNA (*Mus musculas* strain C57BL/6) using primers MCF12 F: 5′-ggctccaggcttgagcatatgtactaatctg-3′ and MCF12 R: 5′-caagaaattggcttcatatgtgatcatcaga-3′. Mouse CFTR exon 12 modifications were also performed by PCR-directed mutagenesis with specific primers.

CFTR exon 12 sequences of different animals were recovered using NCBI BLASTN search. Accession numbers of the sequences are Human (Gene Bank accession number NM_000492.3 *Homo sapiens*), Guinea pig (Gene Bank accession number AF133216.1 *Cavia porcellus*), Ground Squirrel (Gene Bank accession number AC184040.3 *Spermophilus tridecemlineatus*), Mouse (Gene Bank accession number NM_021050.2 *Mus musculus*), Rabbit (*Oryctolagus cuniculus*) GenBank accession no. NM_001082716, Cow (*Bos tortus*) GenBank accession no. NM_174018, Pig (*Sus scrofa*) GenBank accession no. AY585334, Horse (*Equus caballus*) GenBank accession no. NM_001110510, Lemur (*Lemur catta*) GenBank accession no. AC123543.

### Cell culture, transfections and RT–PCR analysis

HeLa cells were cultured in Dulbecco's modified Eagle's medium with Glutamax (Invitrogen) in standard conditions. The minigenes used for transfection were purified using phenol–choloroform extraction followed by a sephacryl S-400 (GE healthcare) column purification step. HeLa cells were plated at a concentration of $2.8 \times 10^5$ to achieve 80–90% confluence. The following day, 500 ng of plasmid DNA were transfected in the cells using Effectene transfection reagents (Qiagen). In case of *in vivo* overexpression of SR proteins (a kind gift from J. Caceres), 1μg of expression vector was mixed with 500 ng of minigenes. Finally, after 24 h total RNA was extracted using TRIreagent solution (Ambion). One microgram of total RNA was used in the retrotranscription reaction with random primers and Moloney murine leukemia virus enzyme (Invitrogen). Spliced products from the transfected minigene were obtained using primers Bra2 5′-taggatccggtcaccaggaagttggttaaatca-3′ and α 2–3 5′-caacttcaagctcctaagccactgc-3′. PCR conditions were the following: 94°C for 5 min.; 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s for 30 cycles; and 72°C for 7 min for the final extension. PCRs were optimized to be in the exponential phase of amplification and products were routinely fractionated in 1.8% (wt/vol) agarose gels.

### *In vitro* transcription and synthetic RNA oligo

RNA was transcribed from PCR templates amplified from the respective plasmids. A T7 promoter sequence was added towards the 5′-end of the template using primer carrying T7 sequence and, similarly, a (TG)8 repeated sequence was used to tag the 3′-end. Every time, 5 μg of DNA template was used in a 60 μl T7 polymerase (Stratagene) transcription reaction. The RNA was then purified using standard Acid–Phenol purification method followed by Ethanol precipitation. Synthetic RNA oligos corresponding to Mouse 17–38, Human 17–38, Mouse 63–81 and Human 63–81 sequences were purchased from Eurofins MWG Operon, Germany.

### Affinity purification of RNA binding proteins and western blot analysis

Ten micrograms of synthetic RNA oligo or 15 μg of transcribed RNA were oxidized in the dark for an hour with sodium *m*-periodate in a 400 μl reaction mixture (100 mM Sodium acetate pH 5.2 and 5 mM Sodium m-periodate). RNA was then ethanol precipitated and resuspended in 200 μl of 100 mM sodium acetate. Approximately 350 μl of prewashed equilibrated adipic acid dehydrazide-agarose beads (50% slurry; Sigma) were added to each oxidized RNA volume and placed in the rotor at 4°C for overnight incubation. This RNA-Bead covalent link formation was also performed in the dark. The immobilized RNA were then washed once with 1 ml of 2 M NaCl and twice using washing buffer (5.2 mM HEPES pH 7.5, 1 mM $MgCl_2$, 0.8 mM Mg acetate). Meanwhile, 200 μl of nuclear extract was mixed with 900 μl RNAse free water, 1× binding buffer (5.2 mM HEPES pH 7.9, 1 mM $MgCl_2$, 0.8 mM Mg acetate, 0.52 mM dithiothreitol, 3.8% glycerol, 0.75 mM ATP, 1 mM GTP and Heparin at the final concentration of 0.5 μg/μl). The RNA-bound beads were then equilibrated in 300 μl of NE mix and incubated for 25 min on a rotor at room temperature. Beads were then washed four times with 1.5 ml washing buffer. In every washing step beads were gently precipitated by gravity on ice. Finally, 50 μl of 3× SDS loading buffer was added and samples were heated for 5 min before loading on a 10% SDS–polyacrylamide denaturing gel. The gel was then electroblotted onto a polyvinylidene difluoride membrane according to standard protocols (Amersham Biosciences) and blocked with 10% skimmed milk (Non fat dry milk in 1× PBS). Membranes targeted for SR protein recognition were blocked using Western blocking reagent (Roche). Proteins were probed with different antibodies and detected with a chemiluminescence kit (ECL; Pierce Biotechnology). Antibodies against hnRNP U and Tra2 β were kind gifts from G. Dreyfuss and I.C Eperon, respectively. Monoclonal Anti-ASF/SF2, SC35 and 1H4 (against SRp 75, 55, 40) antibodies were purchased from Zymed Laboratories Inc. Changes of protein binding levels have been quantified relative to TDP-43 using an Ultro Scan XL, Pharmacia LKB—laser densitometer at 633 nM wavelength according to manufacturer's instructions.

### siRNA knockdown of splicing factors

siRNA sense sequences used for silencing the different target proteins were the following: human hnRNPA1—cagcugaggaagcucuuca (Sigma), and human hnRNPA2—ggaacaguuccguaagcuc (Sigma); human hnRNP C1/C2-gcaaacaagcaguagagau (Sigma), human DAZAP1-gagacucugcgcagcuacu (Dharmacon) and luciferase no. 2 gene control, gccauucuauccucuagaggaug (Dharmacon). HeLa cells were plated at $0.7 \times 10^5$ cells per well in 35-mm plates to achieve 30–40% confluence. The next day, 3 μl Oligofectamine (Invitrogen) was combined with 15 μl of Opti-MEM medium (Invitrogen) and 3 μl of 40 μM siRNA duplex oligonucleotides was diluted in a final volume of 180 μl of Opti-MEM medium. The two mixtures were combined and left for 20 min at room temperature. Finally, this mixture was added to the cells, which were mantained in 0.9 ml of Opti-MEM only. After 6 h, 500 μl of 30% FBS (Foetal bovine serum, Invitrogen) was added. Six to eight hours later Opti-MEM was exchanged with Dulbecco's modified Eagle medium and the cells were transfected with the minigene of interest (500 ng) using Qiagen Effectene transfection reagents. On the third day, HeLa cells were harvested for protein and RNA extractions. RT–PCR from total RNA was performed as for the transfection protocol described above. Whole-protein extracts were obtained by cell sonication in lysis buffer (1× PBS and 1× Protease inhibitor cocktail) and analyzed for hnRNPA1, A2, C1/C2 and DAZAP1 endogenous protein expression by immunoblotting using the antibodies described above. Tubulin was used as total protein loading control.

## RESULTS

In order to better characterize the CERES2 element in CFTR exon 12 we selected two pathological missense mutations (G48C, A51T) and three samesense substitutions (T40C, A49G and C52T) that were already known to affect CFTR exon 12 splicing when inserted in the pTB minigene (Figure 1). In particular, transfection in HeLa cells of the G48C and A49G substitutions caused exon skipping in the full exon 12 context whilst A51T caused full exon inclusion, as previously reported (44). In addition, to widen the number of mutations under study we also chose the T40C and C52T synonymous substitutions that were already known to cause total exon skipping when inserted in the human context (40). These substitutions are particularly interesting form an evolutionary point of view as they are naturally present in the mouse CFTR exon 12 sequence and C52T has also been reported as a human polymorphism/possible mutation in the Cystic Fibrosis Mutation Database (www.genet.sickkids.on.ca). The position and consequences of all these substitutions when inserted in a CFTR exon 12 minigene are summarized in Figure 1A and B, respectively.

First of all, we were interested to see whether the functional effects of these substitutions were dependent on the context provided by the rest of the exon sequence. To study this, we have analyzed their functional effect in

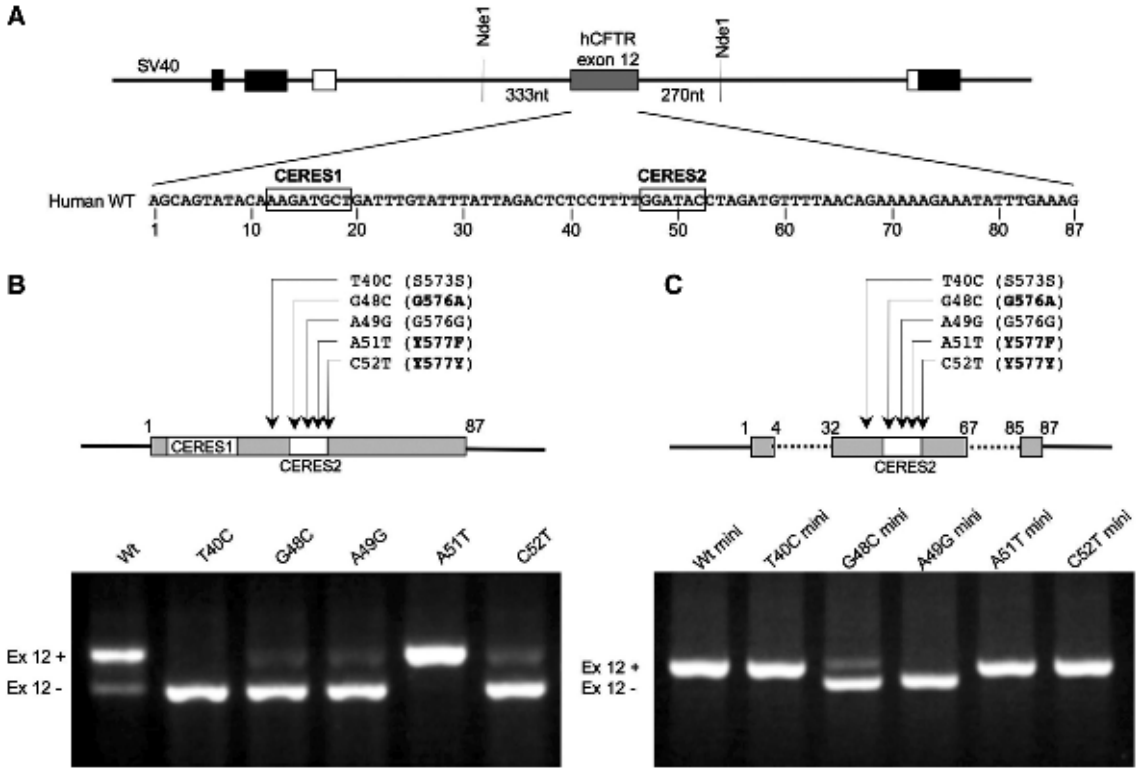


**Figure 1.** (**A**) Schematic presentation of the hybrid minigene used in transfection experiments. The α-globin, fibronectin EDB, and human CFTR exon 12 are shown as black, white and gray boxes, respectively. The sequence of CFTR exon 12 and position of the CERES1 and CERES2 elements is reported in full. (**B** and **C**) Schematic diagram of both full and mini CFTR Ex. 12 constructs used in the analysis. The vertical superimposed arrows indicate the locations of both natural and synonymous mutations. The amplified RT–PCR products stained with ethidium bromide are shown in the bottom panels. Spliced transcripts are shown with Ex.12+ for inclusion and Ex.12− for exclusion of the exon.

a shortened CFTR exon 12 sequence obtained by removing the regions near the 3 and 5′s and downstream regions but maintaining 4 and 3 nts close to the 3′ intron-exon and exon-5′ intron junctions, respectively. We called this construct 'mini' exon 12 (Figure 1C, upper panel). When all the mutations analyzed in Figure 1B were inserted in this reduced context both the G48C and A49G were still capable of inducing exon skipping as observed in the full length exon 12, whilst the enhancing effect of A51T could not be observed owing to the fact that the wild-type mini exon 12 is fully included in the spliced transcript (as opposed to only 80% inclusion of the full length exon 12) (Figure 1C, lower panel). Interestingly, both mouse-specific T40C and C52T substitutions lost the ability to induce exon skipping, suggesting that their effect on the CFTR exon 12 splicing process necessitated the presence of either one or both human flanking regions (Figure 1C, lower panel and Figure 6).

### Identifying the trans-acting factors whose binding is affected by these substitutions

Considering that overexpression of the classical splicing factors hnRNP A1 and SF2/ASF was already described to affect CFTR exon 12 splicing (44) it was decided to better characterize the effect of these substitutions in terms of binding to a wide range of SR and hnRNP splicing factors. To achieve this, we have used a pulldown system previously used in our lab to identify specific RNA binding proteins in a variety of exonic/intronic contexts (46,47). The transcribed RNAs carry a $(UG)_8$ tail that functions as a loading control for the TDP-43 protein (48) (Figure 2A).

In the first analysis, we tested the mini wild-type CFTR exon 12 sequence and two versions carrying the two missense mutations G48C and A51T for binding to the following proteins: hnRNP U, PTB, hnRNP H, DAZAP1, hnRNP C2, A1, A2 and SRp75, SRp55, SRp40, SF2/ASF and Tra2β. The results of this analysis are reported in Figure 2B. This figure shows that no binding could be observed for the hnRNP H, PTB, SRp75, SRp40, SC35 and Tra2β proteins to the mini-exon 12 sequence, both in its wild-type form and carrying either the G48C or the A51T mutations. On the other hand, some of the proteins tested could bind all these sequences, irrespectively of the presence or absence of mutations (hnRNP U and SRp55). Interestingly, a few displayed a differential binding ability in the wild-type sequence with respect to these mutations. In particular, the most striking change could be observed for the SF2/ASF protein that could better bind the A51T mutant with respect to the wild-type sequence. At the same time, it could also bind less efficiently to the G48C mutant than to the wild-type. Finally, hnRNP C2 could also bind less efficiently to the A51T mutant. It is worth noting that

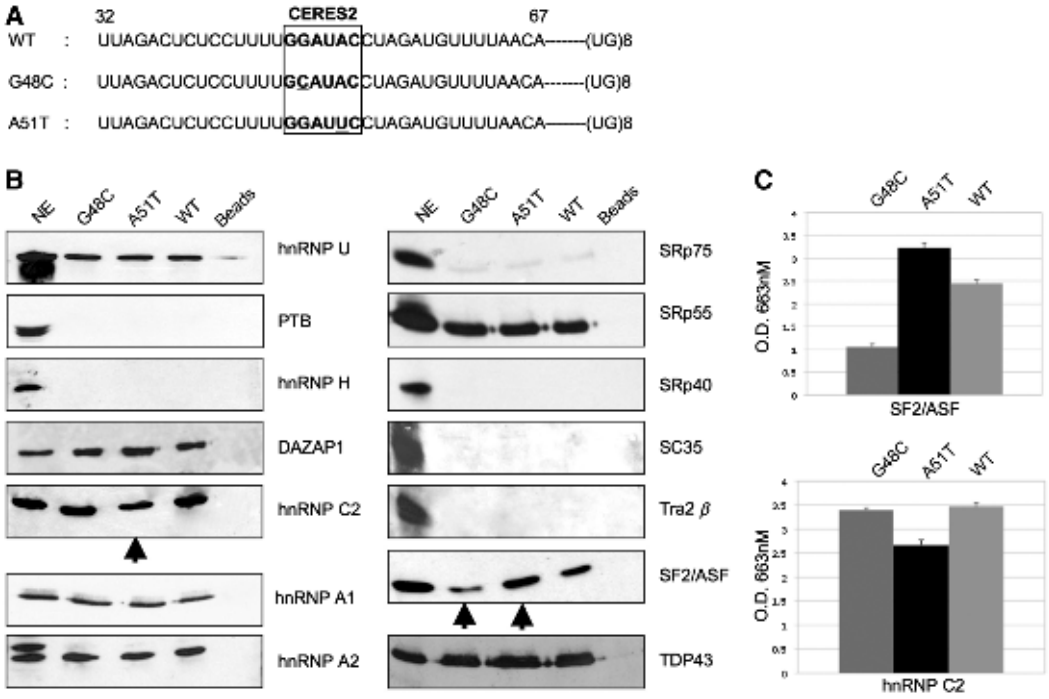


**Figure 2.** Western blot analysis of recovered proteins after pulldown of two naturally occurring nonsense CFTR mutations (G48C/G576A and A51T/Y577F) compared with CFTR Ex. 12 wild type. *In vitro* transcribed RNA was used for analysis. (**A**) Transcribed RNA sequences used for pulldown analysis with the mutations (underlined). A $(UG)_8$ repeat specific for TDP43 was added at the 3′-end of each RNA to normalize the data after western blot. (**B**) Affinity assay for binding of the following hnRNPs: U, PTB, H, DAZAP1, C2, A1 and A2. (**C**) Affinity assay for binding of the following SR factors: SRp75, SRp55, SRp40, SC35, Tra2β and SF2/ASF. Detection of all these proteins was performed by western blot using specific antibodies. Quantification of SF2/ASF and hnRNP C2 binding levels as determined by densitometric analysis are reported in Figure 2C (normalized by TDP-43). Standard deviation values from three independent experiments are shown.

these changes are not all or nothing effects, but they are small changes then amplified by the combinatorial effect of all the other elements involved. On the other hand, the A51T mutant that in the whole exon 12 context has an exon inclusion enhancing effect, displayed increased SF2/ASF binding levels than the wild-type sequence. Quantification of hnRNP C2 and SF2/ASF binding levels in these experiments (normalized against TDP-43) are reported in Figure 2C as determined using densitometric analysis from three independent experiments.

In the case of the three synonymous mutations (T40C, A49G and C52T), the pulldown experiments yielded less varied results (Figure 3B). In fact, no changes could be observed in the binding profiles of RNAs carrying the T40C and C52T substitutions with respect to the wild-type sequence. However, in the case of the A49G mutation we observed less binding of the SF2/ASF protein than in the wild-type sequence and increased binding of hnRNP C2 (quantification of these proteins are reported in Figure 3C), a situation that made the effects of this mutation very similar to those observed for G48C (Figure 2B). Furthermore, it was also consistent with its inhibitory effect in the mini-exon 12 minigene (Figure 1C). In this respect, the observation that no changes could be observed for any of these proteins in the case of T40C and C52T was also consistent with the functional assays demonstrating that these two substitutions were neutral in the human mini-context (Figure 1C).

### Validating the role played by SR factors in CFTR exon 12 splicing

In order to validate the role played by the SR proteins, we tested the response of both the G48C and A49G minigenes to overexpression of the specific interactors of the mini-exon 12 sequence, SF2/ASF and SRp55 and of SC35 (as an example of a SR protein not interacting with the mini exon). The results shown in Figure 4A demonstrate that both SF2/ASF and SRp55 consistently have a higher enhancing effect on the mini exon 12 inclusion levels than SC35, suggesting that direct interaction provides an advantage over the well known generalized enhancing effect of SR proteins. Interestingly, however, the enhancement observed for the two mutants was not the same, with A49G being less responsive especially for SRp55 overexpression than G48C. Finally, it should be noted that deletion of the central CERES2 region also abolished completely the response of the mini-exon 12 to all SR protein overexpression, demonstrating that their action in the mini-exon context is mediated only through the CERES2 sequence. In parallel, to further rule out non-specific effects of SF2/ASF overexpression, we also performed overexpression analysis of a series of deletion mutants (Figure 4B). Also in this case, mutants lacking either the RRM2 region (ΔRRM2) or the RS domain

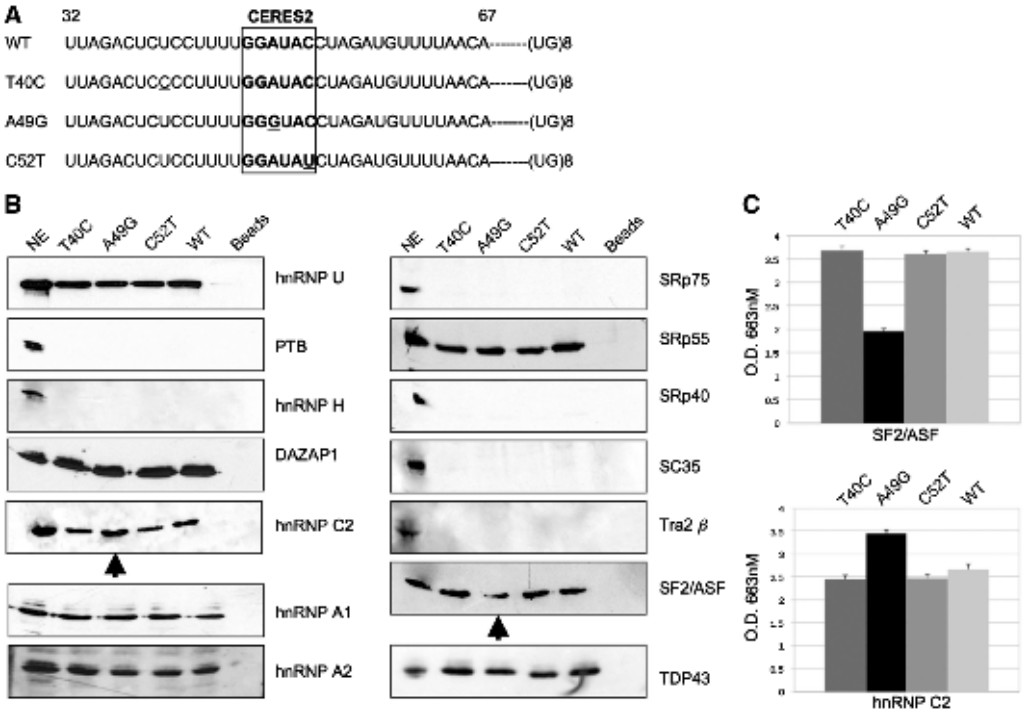


**Figure 3.** Affinity pulldown performed for the three synonymous mutations (T40C, A49G and C52T) in CFTR exon 12. (**A**) Transcribed RNA sequences used for pulldown analysis along with their mutations (underlined). A (UG)8 repeat specific for TDP43 was added at the 3′-end of each RNA for pulldown normalization. (**B**) Affinity pulldown assay for hnRNPs U, PTB, H, DAZAP1, C2, A1 and A2. (**C**) Affinity pulldown assay for SR proteins SRp75, SRp55, SRp40, SC35, Tra2 β and SF2/ASF. Detection of all these proteins was performed by western blot using specific antibodies. Quantification of SF2/ASF and hnRNP C2 binding levels as determined by densitometric analysis is reported in Figure 3C (normalized by TDP-43). Standard deviation values from three independent experiments are shown.

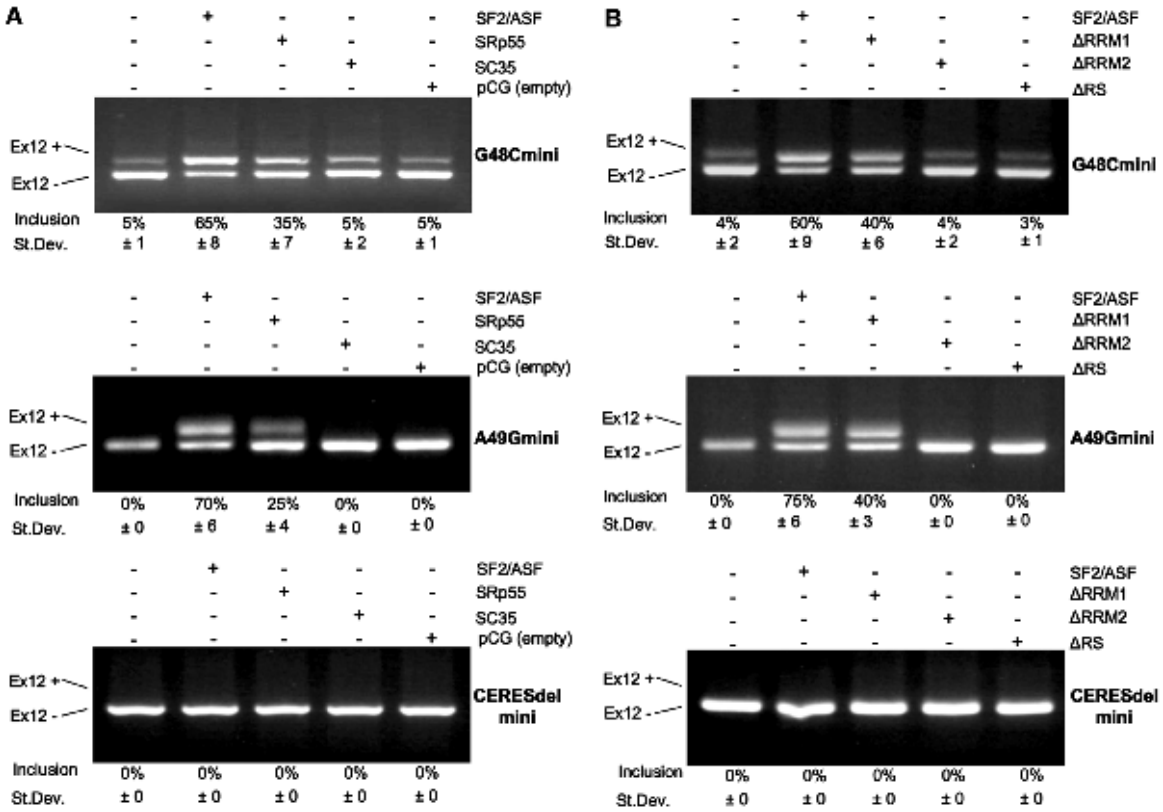


**Figure 4.** (**A**) Analysis showing the *in vivo* overexpression of SR proteins (SRp55, SC35 and SF2/ASF) to rescue CFTR exon 12 mutant minigenes in their mini context. In the case of the CERESdel mini construct, 6 nt of CERES2 (GGATAC) was removed. (**B**) Overexpression of wild-type SF2/ASF and a series of SF2/ASF mutants (ΔRRM1, ΔRRM2 and ΔRS) in the presence CFTR exon 12 mutant minigenes in their mini context. The amplified RT–PCR product of the spliced/unspliced mRNAs are stained using ethidium bromide and run in an agarose 1.8% gel. Exon inclusion and skipping are is shown by Ex12+ and Ex12−, respectively.

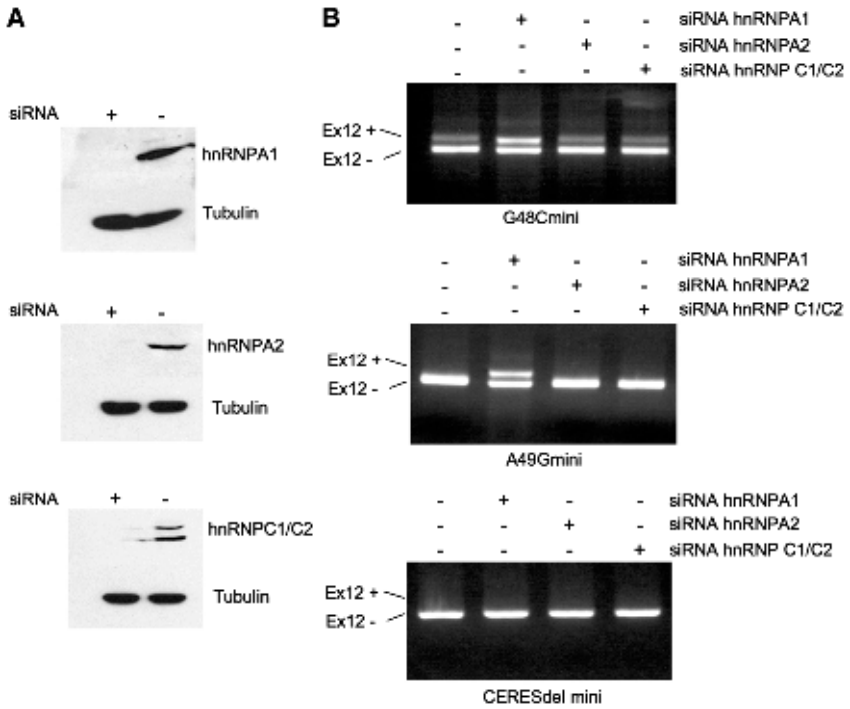


**Figure 5.** (**A**) Knockdown of hnRNP factors in HeLa cells using siRNAs (+) against hnRNP A1, A2 and C2. Luciferase siRNA treated cells have been used as a control and correct depletion was determined by western blot assay (extracts have been normalized using tubulin) (**B**) Effects of transfecting the G48C, A49G and CERESdel minigenes in HeLa cells selectively depleted of each of these hnRNP proteins. RT-PCR samples are stained with ethidium bromide and run on a 1.8% agarose gel. Exon inclusion and skipping are shown by Ex12+ and Ex12−, respectively.

(ΔRS) could not enhance inclusion, further supporting the specificity of this SR protein functional interaction with the CERES2 sequence.

### Validating the role played by hnRNP factors in CFTR exon 12 splicing

Because of their abundance in the nuclear extract, overexpression studies for the different hnRNP proteins did not yield satisfactory results (data not shown). For this reason, in order to test effectively the functional effects of the hnRNP interactors found in pulldown analysis we performed individual siRNA-mediated knockdown of the well known hnRNP A1, A2 and C2 proteins (Figure 5A). As shown in Figure 5B, the only siRNA knockdown that could rescue both the G48C and A49G mini-exons inclusion was hnRNP A1. Importantly, knockdown of this protein had no effect on the CERESdel minigene. Finally, no effect could be detected following hnRNP C2 knockdown which is rather surprising considering that the role of this protein in the regulation of splicing control has been recently re-evaluated in high-throughput studies (23). In addition, as there are many hnRNPs with redundant functions we have also tried the simultaneous depletion of A1, A2 and C2 in different combinations but could not confirm their role in the splicing regulation of CFTR exon 12 (data not shown). It should be noted, however, that these results do not mean that only hnRNP A1 can modulate CFTR exon 12 splicing. In fact, some of these proteins could still play an active role in the presence of reduced amounts of positive factors (i.e. SF2/ASF or SRp55) and further work will be needed to clarify this issue.

### Recovery of the T40C and C52T inhibitory action through the add back of human and mouse flanking CFTR exon 12 sequences

In order to better characterize the mode of action of these substitutions to the mini-exon 12 minigene, we selectively added back the missing upstream and downstream sequences (both in their human and mouse forms) and observed which mutant was able to recover the inhibitory effect of the T40C/C52T substitutions (Figure 6A). As shown in Figure 6B, right panel, the wild type exon 12 constructs that contain the added-back human and mouse upstream regions display full inclusion (constructs A and D). However, when we inserted back the T40C and C52T mutations the inhibitory effect could be detected only in the constructs with the added back human sequence but not with those with the mouse sequence (compare constructs B–C with E–F). A similar situation could also be observed with the downstream regions. In fact, Figure 6B, middle panel, shows that wild-type exon 12 sequences with the added-back human and mouse downstream region display full inclusion (constructs G and J). However, when T40C and C52T were inserted back it was observed that the inhibitory effect could be detected only in the constructs with the human but not with the mouse downstream sequence (compare constructs H–I with K–L). Interestingly, the integrity of maintaining the mouse polypurinic GAAGA ACAAG motif present in the mouse sequence (underlined

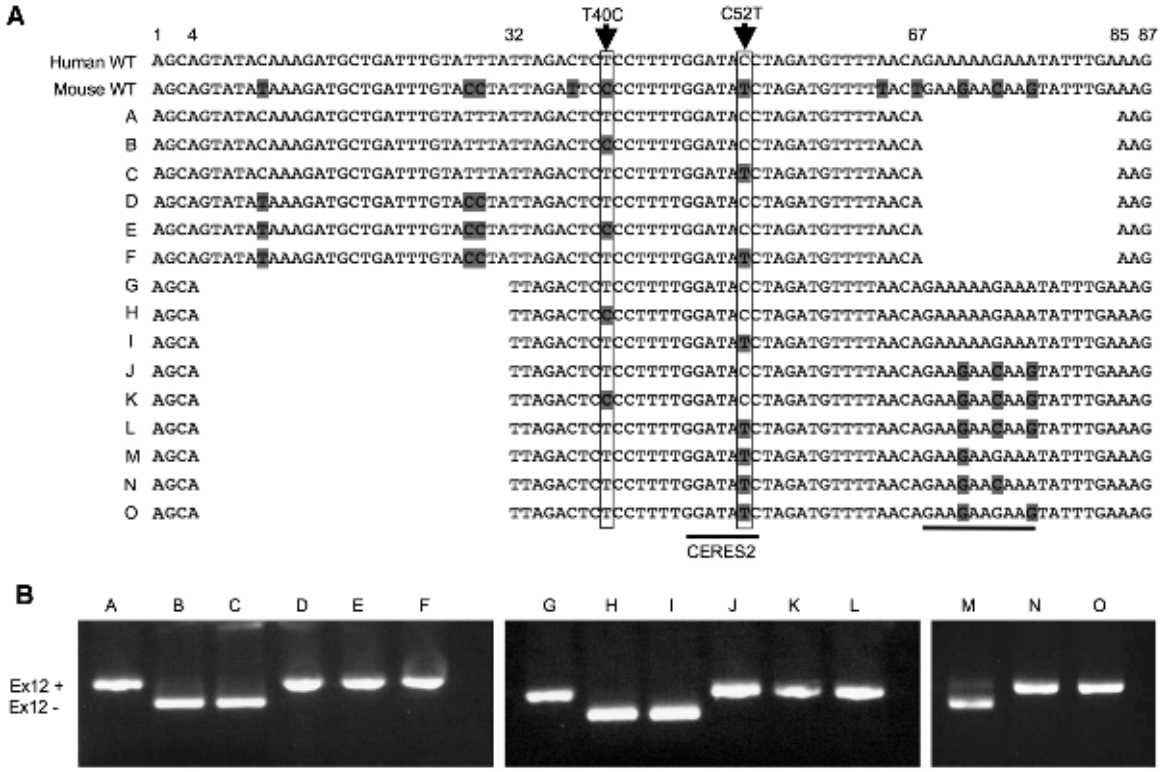


**Figure 6.** (**A**) Schematic diagram of the wild-type and hybrid mouse and human CFTR exon 12 sequences used to construct a series of minigenes (labeled A to O). Nucleotide differences in the mouse sequence with respect to the human are boxed. Blank spaces in the alignment represent the sequences removed form the exon. (**B**) Results of the transfection analysis of the minigenes labeled A to O following transfection in HeLa cells. Exon inclusion is shown by Ex12+ and skipping Ex12−. RT–PCR samples are stained with ethidium bromide and run on a 1.8% agarose gel.

in Figure 6A, bottom) is particularly important. In fact, the majority of substitutions that tend to restore the mouse sequence can successfully withstand the inhibitory action of the C52T substitution (construct N-O as opposed to M).

Taken together, these results suggest that the human and mouse flanking regions have different splicing regulatory properties: human sequences, both upstream and downstream of the central exon 12 region containing CERES2, may be predominantly inhibitory. On the other hand, the mouse upstream and downstream sequences seem to enhance exon recognition.

The hypothesis that mouse sequences enhanced exon recognition was thus tested at the functional level by amplifying a mouse CFTR exon 12 sequence and inserting it in the pTB minigene system (Figure 7A). In this sequence, we then deleted the upstream and downstream regions, either separately or in combination (mutants A–C). The results of this analysis are reported in Figure 7B and show that deleting only the upstream sequence (mutant A) had no effect of mouse CFTR exon 12 inclusion levels. On the other hand, deletion of the downstream sequence (mutant B) resulted in ~15% exon skipping. Interestingly, if both regions are deleted at the same time (mutant C) the levels of exon skipping increase to 25%, indicating that also the polypurinic downstream sequence can function as an ESE once the upstream ESE sequence is absent.

### *Trans*-acting factors binding to the human and mouse 17–38 and 63–81 sequences

Based on these results, it was thus likely that mouse and human flanking sequences could bind a different set of proteins. For this reason, we selected the 17–38 and 63–81 mouse and human regions (Figure 8A) to perform pulldown assays as previously described for the central region (Figures 2–3). The Western blot analysis to check for SR protein binding showed that both mouse sequences could bind SF2/ASF, SRp75 and SRp55 more efficiently than the respective human sequences. In addition, mouse nucleotide stretch 64–82 is also capable of binding SRp40 whilst the human 64–82 sequence is not (Figure 8C). As control, the recognition with antibodies against hnRNP proteins A1 and U showed that both these factors could bind with equally well with the mouse and human sequences. Unfortunately, using this approach we were unable to determine which factors are responsible for the ESS activity of the human 18–39 and 64–82 sequences. Of course, the most probable reason is that we have tested only a fraction of the candidates and there are many potential known or unknown proteins that could exert an effect on splicing regulation. Further work is currently under way to better define this point. In any case, these results were consistent with the *in vivo* results that detected different functional properties between the mouse and human flanking regions.

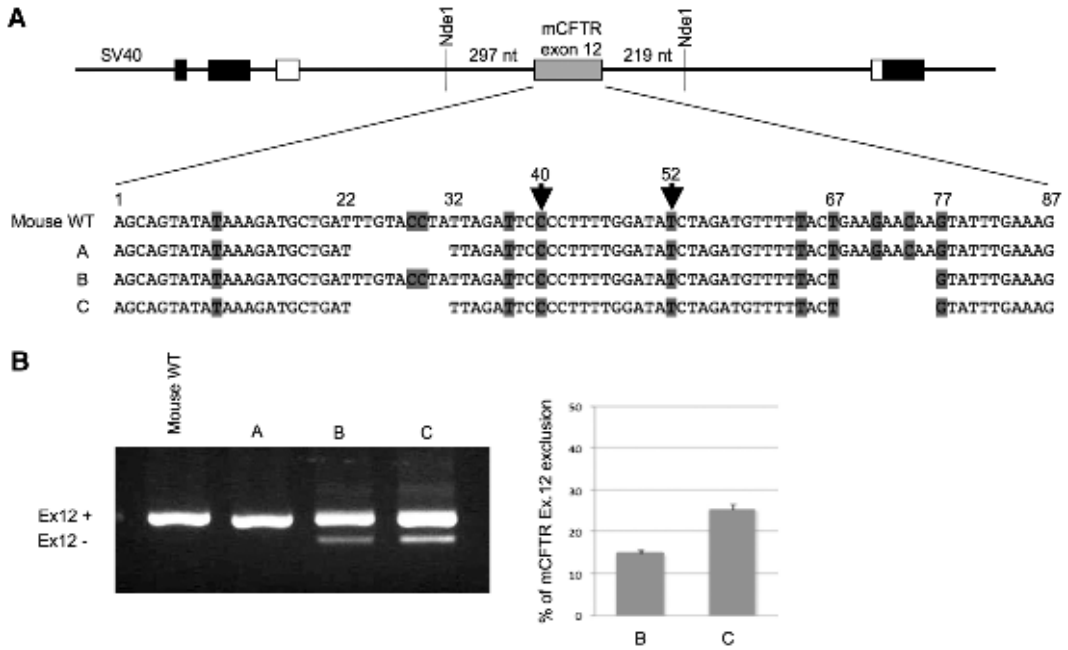


**Figure 7.** (**A**) Schematic analysis of the pTB minigene containing the mouse CFTR exon12 sequence (reported in full). Mouse specific nucleotides are boxed. Mutants A, B and C show the combination of nucleotide deletions spanning from nucleotides 23 to 31 and 68 to 76 included (**B**) Mouse CFTR exon 12 minigenes (wild-type, A, B and C) were transfected in HeLa cells for *in vivo* splicing assay. The amplified RT–PCR products of this splicing assay are stained by ethidium bromide and run on a 1.8% agarose gels. Exon inclusion is shown by Ex12+ and skipping by Ex12−.

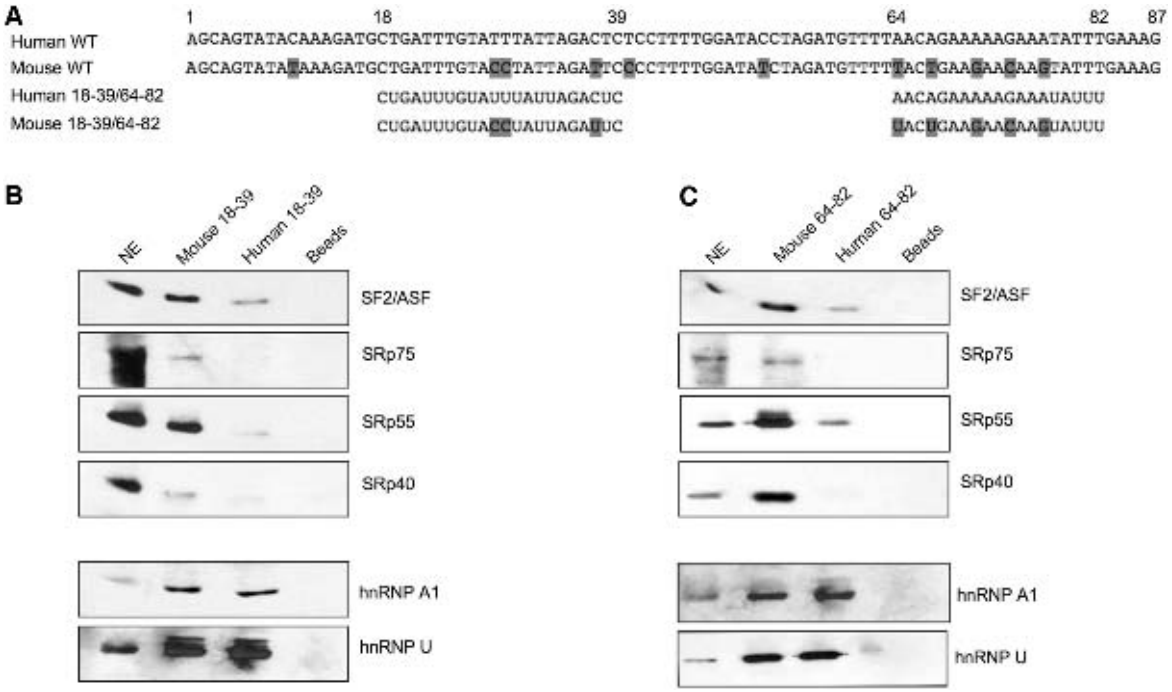


**Figure 8.** (**A**) Mouse and Human synthetic RNA sequences used in a affinity purification assay. Mouse specific nucleotides are boxed. (**B**) Pulldown analysis performed for human and mouse CFTR exon 12 sequences spanning from nucleotide 18–39 were comparatively analyzed for splicing factors. Harvested proteins were targeted for SR proteins (SRp 75, SRp 55, SRp 40, SF2/ASF) in western blot analysis. The membrane was later probed with hnRNP A1 and U antibody. (**C**). Similar comparative analysis for splicing factors was performed for both the human and mouse CFTR exon 12 sequences spanning nucleotides 64–82.

## DISCUSSION

Missense mutations in human CFTR exon 12 have been described to be the causative agent of Cystic Fibrosis through the inactivation of a highly conserved region that encodes part of the first nucleotide binding fold of the protein (49). In particular, among all disease-causing mutations known to affect this exon, several have been described to induce its skipping during the splicing

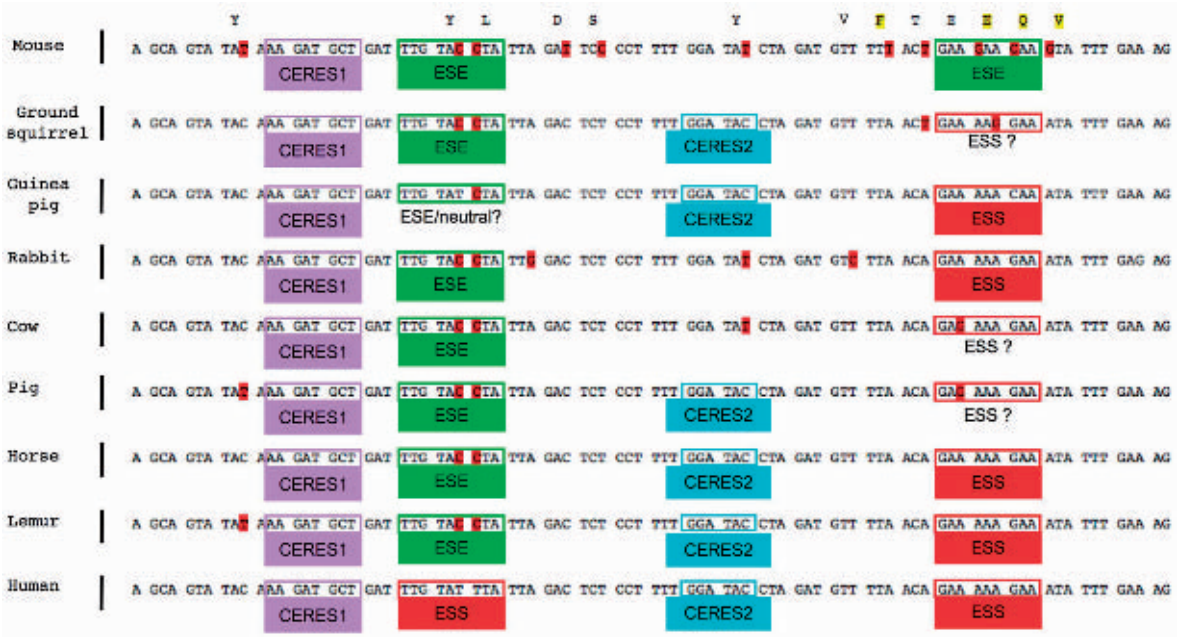


**Figure 9.** Schematic comparison and conservation of the splicing regulatory elements (ESEs, ESSs, and CERES) in different species based on our functional analysis.

process (50–52). In our lab, we originally defined within CFTR exon 12 two regions, named CERES1 and CERES2 (Figure 9), that functioned in a highly context-dependent manner to regulate the splicing process of this exon (44). In fact, the effect of natural and artificial mutations within these regions could not be predicted easily using current bioinformatics approaches (44), highlighting recent recommendations that these programs should be used with caution when they are used as a diagnostic tool (53,54). Indeed, successive experimental comparison between human and mouse CFTR exon 12 sequences demonstrated that about one quarter of all artificial combinations of mouse–human same sense substitutions resulted in exon skipping (40). Taken together, these findings suggested that the whole coding sequence of CFTR exon 12 is under strong selective pressures not only for functional reasons at the protein level, but also for the maintenance of proper exon recognition by the splicing machinery.

Up to now, however, the detailed molecular bases of CERES action were not known. We have performed an analysis of the trans-acting factors that bind the CERES2 element localized in CFTR exon 12. Our analysis was preliminarily focused at characterizing the binding properties of the most common splicing regulators, and in particular those belonging to either the SR or the hnRNP class of *trans*-acting factors. The results of our analysis have demonstrated that in normal conditions human CERES2 can bind a substantial number of these proteins, something that may be rather surprising since the core CERES2 element represents a very short stretch of RNA sequence (<10 nucleotides). Among SR proteins, we have found SF2/ASF and SRp55 whilst regarding hnRNP proteins specific binding could be detected for most hnRNP A/B family members. Most importantly, the relative binding capacity of some of the factors was modified following the introduction of disease-associated missense mutations or of samesense substitutions that were already known to affect CFTR exon 12 inclusion levels. From a basic RNA binding protein point of view, this finding highlights the great flexibility provided by RRM motifs that can recognize a few specific bases at selected positions using their main chains and then use side-chain interactions to stabilize binding (55). This probably explains why the CERES2 sequence rather than functioning as a binding site for a single protein only can function as a kind of aggregation site for many SR/hnRNP factors. Evidence of a functional interaction was confirmed only for SF2/ASF, SRp55 and hnRNP A1. However, it should be noted that these experiments were performed in a severely reduced context (mini-exon 12) and that in a more natural setting many of these proteins will be able to play a role, especially considering the fact that the exonic flanking regions also can bind several SR/hnRNP factors (see below). At the moment, of course, we have tested only few regulatory proteins and hence cannot rule out the presence of additional yet unknown factors that might also contribute to define/hinder CFTR exon 12. Another possibility is that our substitutions may affect the RNA secondary structure of CFTR exon 12. However, an evolutionary-based model of CFTR exon 12 RNA secondary structure has already been previously published by Meyer and Miklos (56). In their work, they have also analyzed the potential impact of some splicing mutations including substitutions in the 40T and 52C positions. The conclusion is that there are only marginally significant changes in the RNA structure of the mutants with respect to the predicted wild-type sequence. It is for this reason that we decided to concentrate on analyzing *trans*-acting factors rather than taking

into consideration structural changes following our substitutions.

Taken together, our results suggest that this crowding together of many different proteins (both positive and negative in terms of their effect on splicing) may explain why single-point substitutions within the CERES2 element have such an unpredictable effect on the exon recognition levels. Up to now, on a slightly wider scale the CERES sequence situation is similar to what has already been found in several small exons, such as SMN exon 7, where *trans*-acting factors (SF2/ASF and hnRNP A1) binding to the same exonic region contribute to exon inclusion/skipping (57,58). Other examples of very complex systems include the human *c-src* exon N1 (59,60), *CD44* exon v5 (61–63), *HipK3* 'T' exon (64), and chicken *cTNT* exon 5 (65) where numerous factors have been shown to participate in splicing regulation in close spatial proximity to each other. Indeed, for SMN exon 7 it has been hypothesized the existence of an Extended Inhibitory Context (Exinct) that is caused by overlapping regulatory motifs not all of which have still been characterized in depth (66).

On a more general note, the existence of these numerous splicing factor binding sites co-existing together on the same exon has very important implications with regards to evolutionary constraints in codon composition. For example, our results have shown very clearly that the inhibitory effect of some synonymous nucleotide substitutions (A40T and C52T) naturally occurring in the mouse sequence can be explained by the different splicing regulatory properties of the human and mouse flanking exonic sequences (summarized in Figure 9). In particular, by comparing the ESE, ESS, and CERES elements several considerations can be made with regards to the sequence changes that have occurred in the mouse to human transition during the course of evolution. In fact, as shown in Figure 9 it can be hypothesized that the creation of the human CERES2 element in CFTR exon 12 has relieved the pressure to keep the two weak ESE elements loosely localized in the mouse 18–39 and 64–82 flanking regions. Only after the creation of CERES2 element these two regions could thus undergo nucleotide substitutions that either weakened these elements (i.e. in the Ground squirrel and Guinea pig 64–82 region) or even changed them to functional silencer sequences (in human 18–39 and 64–82 region). The advantage of these coding region changes can be only hypothesized at this stage, but it may involve further steps beyond mRNA processing such as enzymatic activity or protein stability. Irrespective of their significance, however, the important issue is that sequence changes could only be introduced through the creation of the CERES2 element (the importance of which is highlighted by the observation that many of the mouse to human substitutions analyzed in this study lead directly to total exon skipping). It is also interesting to note that a comparison of these sequence elements also in other species (Figure 9) does not contradict the conclusions we have drawn from mouse versus humans. In fact, for example, both cows and rabbits that do not contain the CERES2 element have absolutely conserved ESE sequences. Of course, additional experiments will need to be performed before we can draw firm conclusions on this issue. Nonetheless, in our work, we show that an integrated analysis of *cis*- and *trans*-acting factors binding to exonic elements can provide a substantial wealth of information on potential evolutionary mechanisms.

Looking at the splicing regulatory elements in Figure 9 it is also possible to draw some additional conclusions with regards to our understanding of splicing regulation in general. It is clear that, the CFTR exon 12 sequence is literally covered by regulatory elements that we probably still consider (rather mistakenly) as separate elements. Proof of this is the observation that the activity of many of these splicing regulatory elements (especially CERES) cannot be exported in different contexts (44). Indeed, our results point towards a situation where in exons like CFTR exon 12 we should virtually consider every nucleotide as potentially capable of affecting splicing inclusion levels. The only question that might then remain to us is the direction of this change (whether increased inclusion/skipping) and its extent. It is very probable that as our knowledge of splicing systems increases this kind of situations will be more and more common. From a practical point of view, this will have several consequences. From a clinical point of view, increased importance have to be given at analyzing RNA transcripts directly from patient tissues or, routinely, through minigene based systems that will mimic this kind of global splicing regulatory networks (67). Second, this increased awareness will be useful for the development of novel bioinformatics methods aimed at predicting splicing outcomes that, until now, have been primarily focused at considering enhancer and silencer elements as well distinct entities with rather limited success (53,54). Finally, it should gradually shift our view of splicing where exon inclusion levels should not always be viewed as the straightforward algebraic sum of enhancer/silencer elements but as rather as an integrated unit, where silencing and enhancing functions may functionally overlap to a degree that has often been underestimated.

## FUNDING

Telethon Onlus Foundation (grant no. GGP06147 to F.E.B.); European community grant (EURASNET-LSHG-CT-2005-518238 to F.E.B.). Funding for open access charge: ICGEB core funding.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Pagani,F. and Baralle,F.E. (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat. Rev. Genet.*, **5**, 389–396.
2. Parmley,J.L. and Hurst,L.D. (2007) How do synonymous mutations affect fitness? *Bioessays*, **29**, 515–519.
3. Chamary,J.V., Parmley,J.L. and Hurst,L.D. (2006) Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat. Rev. Genet.*, **7**, 98–108.

4. Black,D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.*, **72**, 291–336.
5. Mardon,H.J., Sebastio,G. and Baralle,F.E. (1987) A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic Acids Res.*, **15**, 7725–7733.
6. Reed,R. and Maniatis,T. (1986) A role for exon sequences and splice-site proximity in splice-site selection. *Cell*, **46**, 681–690.
7. Breathnach,R., Benoist,C., O'Hare,K., Gannon,F. and Chambon,P. (1978) Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc. Natl Acad. Sci. USA*, **75**, 4853–4857.
8. Ben-Dov,C., Hartmann,B., Lundgren,J. and Valcarcel,J. (2008) Genome-wide analysis of alternative pre-mRNA splicing. *J. Biol. Chem.*, **283**, 1229–1233.
9. Moore,M.J. and Silver,P.A. (2008) Global analysis of mRNA splicing. *RNA*, **14**, 197–203.
10. Buratti,E. and Baralle,F.E. (2005) Another step forward for SELEXive splicing. *Trends Mol. Med.*, **11**, 5–9.
11. Chasin,L.A. (2007) Searching for splicing motifs. *Adv. Exp. Med. Biol.*, **623**, 85–106.
12. Wang,J., Smith,P.J., Krainer,A.R. and Zhang,M.Q. (2005) Distribution of SR protein exonic splicing enhancer motifs in human protein-coding genes. *Nucleic Acids Res.*, **33**, 5053–5062.
13. Pozzoli,U. and Sironi,M. (2005) Silencers regulate both constitutive and alternative splicing events in mammals. *Cell. Mol. Life Sci.*, **62**, 1579–1604.
14. Cartegni,L., Chew,S.L. and Krainer,A.R. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.*, **3**, 285–298.
15. Blencowe,B.J. (2000) Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.*, **25**, 106–110.
16. Venables,J.P. (2007) Downstream intronic splicing enhancers. *FEBS Lett.*, **581**, 4127–4131.
17. Wang,Z., Rolish,M.E., Yeo,G., Tung,V., Mawson,M. and Burge,C.B. (2004) Systematic identification and analysis of exonic splicing silencers. *Cell.*, **119**, 831–845.
18. Wang,Z. and Burge,C.B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA*, **14**, 802–813.
19. Lin,S. and Fu,X.D. (2007) SR proteins and related factors in alternative splicing. *Adv. Exp. Med. Biol.*, **623**, 107–122.
20. Huang,Y. and Steitz,J.A. (2005) SRprises along a messenger's journey. *Mol. Cell.*, **17**, 613–615.
21. Long,J.C. and Caceres,J.F. (2009) The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.*, **417**, 15–27.
22. Martinez-Contreras,R., Cloutier,P., Shkreta,L., Fisette,J.F., Revil,T. and Chabot,B. (2007) hnRNP proteins and splicing control. *Adv. Exp. Med. Biol.*, **623**, 123–147.
23. Venables,J.P., Koh,C.S., Froehlich,U., Lapointe,E., Couture,S., Inkel,L., Bramard,A., Paquet,E.R., Watier,V., Durand,M. *et al.* (2008) Multiple and specific mRNA processing targets for the major human hnRNP proteins. *Mol. Cell. Biol.*, **28**, 6033–6043.
24. Jurica,M.S. and Moore,M.J. (2003) Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell.*, **12**, 5–14.
25. Zhu,J., Mayeda,A. and Krainer,A.R. (2001) Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell.*, **8**, 1351–1361.
26. Eperon,I.C., Makarova,O.V., Mayeda,A., Munroe,S.H., Caceres,J.F., Hayward,D.G. and Krainer,A.R. (2000) Selection of alternative 5′ splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Mol. Cell. Biol.*, **20**, 8303–8318.
27. 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.
28. Relogio,A., Ben-Dov,C., Baum,M., Ruggiu,M., Gemund,C., Benes,V., Darnell,R.B. and Valcarcel,J. (2005) Alternative splicing microarrays reveal functional expression of neuron-specific regulators in Hodgkin lymphoma cells. *J. Biol. Chem.*, **280**, 4779–4784.
29. Grosso,A.R., Gomes,A.Q., Barbosa-Morais,N.L., Caldeira,S., Thorne,N.P., Grech,G., von Lindern,M. and Carmo-Fonseca,M. (2008) Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res.*, **36**, 4823–4832.
30. Shin,C., Feng,Y. and Manley,J.L. (2004) Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. *Nature*, **427**, 553–558.
31. Blaustein,M., Pelisch,F. and Srebrow,A. (2007) Signals, pathways and splicing regulation. *Int. J. Biochem. Cell. Biol.*, **39**, 2031–2048.
32. Huang,Y., Yario,T.A. and Steitz,J.A. (2004) A molecular link between SR protein dephosphorylation and mRNA export. *Proc. Natl Acad. Sci. USA*, **101**, 9666–9670.
33. Misteli,T., Caceres,J.F., Clement,J.Q., Krainer,A.R., Wilkinson,M.F. and Spector,D.L. (1998) Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *J. Cell. Biol.*, **143**, 297–307.
34. Buratti,E., Dhir,A., Lewandowska,M.A. and Baralle,F.E. (2007) RNA structure is a key regulatory element in pathological ATM and CFTR pseudoexon inclusion events. *Nucleic Acids Res.*, **35**, 4369–4383.
35. Buratti,E., Muro,A.F., Giombi,M., Gherbassi,D., Iaconcig,A. and Baralle,F.E. (2004) RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. *Mol. Cell. Biol.*, **24**, 1387–1400.
36. Hiller,M., Zhang,Z., Backofen,R. and Stamm,S. (2007) Pre-mRNA secondary structures influence exon recognition. *PLoS Genet.*, **3**, e204.
37. Shepard,P.J. and Hertel,K.J. (2008) Conserved RNA secondary structures promote alternative splicing. *RNA*, **14**, 1463–1469.
38. Cheah,M.T., Wachter,A., Sudarsan,N. and Breaker,R.R. (2007) Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature*, **447**, 497–500.
39. Graveley,B.R. (2005) Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell.*, **123**, 65–73.
40. Pagani,F., Raponi,M. and Baralle,F.E. (2005) Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. *Proc. Natl Acad. Sci. USA*, **102**, 6368–6372.
41. Parmley,J.L., Chamary,J.V. and Hurst,L.D. (2006) Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. *Mol. Biol. Evol.*, **23**, 301–309.
42. Parmley,J.L. and Hurst,L.D. (2007) Exonic splicing regulatory elements skew synonymous codon usage near intron-exon boundaries in mammals. *Mol. Biol. Evol.*, **24**, 1600–1603.
43. Irimia,M., Rukov,J.L. and Roy,S.W. (2009) Evolution of alternative splicing regulation: changes in predicted exonic splicing regulators are not associated with changes in alternative splicing levels in primates. *PLoS ONE*, **4**, e5800.
44. Pagani,F., Stuani,C., Tzetis,M., Kanavakis,E., Efthymiadou,A., Doudounakis,S., Casals,T. and Baralle,F.E. (2003) New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum. Mol. Genet.*, **12**, 1111–1120.
45. Pagani,F., Buratti,E., Stuani,C. and Baralle,F.E. (2003) Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9. *J. Biol. Chem.*, **278**, 26580–26588.
46. Buratti,E., Baralle,M., De Conti,L., Baralle,D., Romano,M., Ayala,Y.M. and Baralle,F.E. (2004) hnRNP H binding at the 5′ splice site correlates with the pathological effect of two intronic mutations in the NF-1 and TSHbeta genes. *Nucleic Acids Res.*, **32**, 4224–4236.
47. 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.
48. Buratti,E. and Baralle,F.E. (2001) 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.*, **276**, 36337–36343.
49. 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.
50. 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.
51. 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.
52. 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.
53. Hartmann,L., Theiss,S., Niederacher,D. and Schaal,H. (2008) Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases? *Front Biosci.*, **13**, 3252–3272.
54. Houdayer,C., Dehainault,C., Mattler,C., Michaux,D., Caux-Moncoutier,V., Pages-Berhouet,S., d'Enghien,C.D., Lauge,A., Castera,L., Gauthier-Villars,M. *et al.* (2008) Evaluation of in silico splice tools for decision-making in molecular diagnosis. *Hum. Mutat.*, **29**, 975–982.
55. Auweter,S.D., Oberstrass,F.C. and Allain,F.H. (2006) Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res.*, **34**, 4943–4959.
56. Meyer,I.M. and Miklos,I. (2005) Statistical evidence for conserved, local secondary structure in the coding regions of eukaryotic mRNAs and pre-mRNAs. *Nucleic Acids Res.*, **33**, 6338–6348.
57. Cartegni,L., Hastings,M.L., Calarco,J.A., de Stanchina,E. and Krainer,A.R. (2006) Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am. J. Hum. Genet.*, **78**, 63–77.
58. Kashima,T., Rao,N., David,C.J. and Manley,J.L. (2007) hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Hum. Mol. Genet.*, **16**, 3149–3159.
59. Sharma,S., Falick,A.M. and Black,D.L. (2005) Polypyrimidine tract binding protein blocks the 5′ splice site-dependent assembly of U2AF and the prespliceosomal E complex. *Mol. Cell.*, **19**, 485–496.
60. Chou,M.Y., Underwood,J.G., Nikolic,J., Luu,M.H. and Black,D.L. (2000) Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing. *Mol. Cell.*, **5**, 949–957.
61. Matter,N., Herrlich,P. and Konig,H. (2002) Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature*, **420**, 691–695.
62. Matter,N., Marx,M., Weg-Remers,S., Ponta,H., Herrlich,P. and Konig,H. (2000) Heterogeneous ribonucleoprotein A1 is part of an exon-specific splice-silencing complex controlled by oncogenic signaling pathways. *J. Biol. Chem.*, **275**, 35353–35360.
63. Cheng,C. and Sharp,P.A. (2006) Regulation of CD44 alternative splicing by SRm160 and its potential role in tumor cell invasion. *Mol. Cell. Biol.*, **26**, 362–370.
64. Venables,J.P., Bourgeois,C.F., Dalgliesh,C., Kister,L., Stevenin,J. and Elliott,D.J. (2005) Up-regulation of the ubiquitous alternative splicing factor Tra2beta causes inclusion of a germ cell-specific exon. *Hum. Mol. Genet.*, **14**, 2289–2303.
65. Ladd,A.N., Stenberg,M.G., Swanson,M.S. and Cooper,T.A. (2005) Dynamic balance between activation and repression regulates pre-mRNA alternative splicing during heart development. *Dev. Dyn.*, **233**, 783–793.
66. Singh,N.N., Androphy,E.J. and Singh,R.N. (2004) An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy. *Biochem. Biophys. Res. Commun.*, **315**, 381–388.
67. Baralle,D., Lucassen,A. and Buratti,E. (2009) Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep.*, **10**, 810–816.