SUPP. MATERIALS AND METHODS

In silico analysis

The free online software packages used in the study are summarized in Supp. Table S1. Local 2D-structures were predicted using RNAshapes 2.1.6 software (available at http://bibiserv.techfak.uni-bielefeld.de/rnashapes/). The minimal free energy of each shape predicted was also calculated (expressed in kcal/mol).

Electrophoretic mobility shift assay

To prepare probes for EMSA, we first performed PCR amplifications as to introduce a T7 promoter adjacent to WT and mutant *CFTR* exon 3 cloned in pET01 plasmids (5ng) using the following primers:

F_CCAAGCTTCTAATACGACTCACTATAGGGAGAAGAATGGGATAGAGAGCTGGC and R_CCCTAAATATAAAAAGATTCCATAGAACA, where the T7 promoter is underlined. PCR products were gel purified and quantified using a Nanodrop. RNA synthesis was performed using the T7 RNA synthesis kit (Roche) using 500ng of T7-labelled CFTR exon 3, 0.5mM of each NTP, 2μ l of T7 polymerase and 20 µCi of CTP(P³³ (PerkinElmer) in a final volume of 40µl. Reaction was incubated at 37°C for 1h30 after which 2µl of DNase were added for an additional 15min (ThermoFisher). Volumes were adjusted to 200µl and RNA probes were phenol extracted and precipitated using ammonium acetate (200µl, 5M), 20mg/ml glycogen and 1ml 100% ethanol. Precipitates were washed with 70% ethanol and air-dried before being resuspended in 100µL water. Increasing quantities of recombinant SF2/ASF (200ng, 400ng and 600ng, from Abnova) were incubated with radiolabeled RNA probes (8µl) for 30 min at 37°C in binding buffer containing 20 mM Hepes (pH 7.4), 1 mM EDTA, 10 mM (NH₄)₂SO₄, 1 mM DTT, 0.2% Tween-20, 30 mM KCl, $7.5\mu g/\mu l$ heparin, and 5% glycerol in a final volume of 20 μl . The probes were denatured at 95°C for 3 min before being cooled on ice and incubated with recombinant SF2/ASF. Samples were separated on a 5% non-denaturing polyacrylamide gel and dried before being scanned using a phospho-imager Storm 860 (Molecular Dynamics).

SUPPLEMENTARY FIGURES





Figure S1: Analyses of constitutive CFTR exons acceptor or donor splice sites.

Analyses were performed using: (A) ESEfinder, FSplice, GeneID, GENSCAN, HBond, HSF and MaxEntScan and (B) NetGene2, NNSplice, SplicePort, SplicePredictor, SpliceView and SSF. Are represented *C190* boxplots of the mean strength of acceptor and donor sites calculated for all *CFTR* exons. Exons with scores in the outliers inferior/equal to the lower inner fence are indicated.

B



Figure S2. C190 boxplots of the mean densities of SREs calculated with the SKIPPY software.

(A) *CI90* box plot of the mean frequency of exon splicing enhancers (ESE) and exon splicing silencers (ESS) localised in the first and last 50bp of each exon, calculated for all *CFTR* exons with the SKIPPY tool (left side) and ESS densities in the intron 100bp upstream or downstream of each splice site junction (right side). Exons presenting with ESS scores and ESS/ESE ratio above the upper inner fence (*CI90*) or with ESE score under the lower inner fence (*CI90*) are indicated.



Figure S3. Hybrid minigene splicing assays for CFTR exon 3 mutational analyses.

Examples of capillary electrophoresis analysis of RT-PCR products obtained after 25 cycles. RNA was purified from BEAS-2B cells transfected with minigene bearing the c.220C>T substitution (A), the c.223C>T substitution (B), the c.224G>T substitution (C), or the c.224G>A substitution (D). RT-PCR was performed using a Fluorescein amidite (FAM)-labeled forward primer located within the splice donor exon and a reverse primer within the splice acceptor exon of the pET01 plasmid. The corresponding size and relative amount for each peak is indicated. The peak at 245 bp corresponds to exon 3-skipped mRNA, whereas the peak at 354 bp corresponds to full-length mRNA.



Figure S4. Cell specific effects an exon 3 splicing.

Exon 3 exclusion represented as mean fold- increase normalized to the corresponding WT control. The cell line tested and the transfected construct are indicated. Experiments were repeated 3 to 5 times for each value. All values are statistically significant (p < 0.01) compared to the corresponding control, except the c.223C>T construct transfected into HeLa cells (#).



Figure S5. SF2/ASF modulation affects mutant c.220C>T exon 3 inclusion.

Upper panel: semi-quantification of exon 3 exclusion measured for WT sequence or for sequences containing either the c.220C>T mutation, in IB3-1 cells. *Lower panel*: Western blot analysis measuring SF2/ASF and LaminB1 expression in corresponding cells. Mean % of control measures of SF2/ASF normalized to LaminB1 signal.



Figure S6. *CI90* **boxplots of the mean size of** *CFTR* **exons.** The sizes of all 27 *CFTR* exons were compared with the median value (*CI90*). Exons with a size in the outliers (upper/lower inner fences, *CI90*) are represented.



Figure S7. In silico analysis of exon 3 structure.

Two-dimensional structures of exon 3 bearing the indicated nucleotide substitution were assessed using RNAshapes 2.1.6. Structures corresponding to nucleotides r.194-247 are illustrated, except for r.220c>u, which corresponds to nucleotides r.185-239. The SF2/ASF-binding motif is indicated, when present, with a black line, and the region of interest (r.217-227) is highlighted with a gray line for structures corresponding to r.224g>a, r.224g>u and r.220c>u. Calculated free energy ranged from -16.3 kcal/mol to -14.7 kcal/mol for WT, from -16.7 kcal/mol to -15.2 kcal/mol for r.224g>a, from -16.5 kcal/mol to -14.9 kcal/mol for r.223c>u, from -16.1 kcal/mol to -14.7 kcal/mol for r.224g>u, and from -16.6 kcal/mol to -15.3 kcal/mol for r.220c>u.

SUPPLEMENTARY TABLES

Name of the software	Link	Reference
ESEfinder3.0	http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi	(Cartegni, et al., 2003)
EX-SKIP	http://ex-skip.img.cas.cz/	(Raponi, et al., 2011)
FSplice	http://linux1.softberry.com/berry.phtml?topic=fsplice&group=p	<u>_</u>
1 Splice	rograms&subgroup=gfind	-
GeneID	http://genome.crg.es/software/geneid/geneid.html	(Blanco, et al., 2007)
GENSCAN	http://genes.mit.edu/GENSCAN.html	(Burge and Karlin. 1997)
HBond	http://www.uni-duesseldorf.de/rna/html/hbond_score.php	-
Human Splicing Finder	http://www.umd.be/HSF/	(Desmet, et al., 2009)
MayEntScon	http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.ht	(Vac and Purge 2004)
WIAXEIIUSCAII	ml	(100 and Burge, 2004)
NetGene2	http://www.cbs.dtu.dk/services/NetGene2/	(Brunak, et al., 1991)
NNSplice	http://www.fruitfly.org/seq_tools/splice.html	(Reese, et al., 1997)
SKIPPY	http://research.nhgri.nih.gov/skippy/input.shtml	(Woolfe, et al., 2010)
SplicePort	http://spliceport.cbcb.umd.edu/SplicingAnalyser.html	(Dogan, et al., 2007)
SplicePredictor	http://spliceport.chch.umd.adu/SplicingAnalyser.html	(Brendel and Kleffe,
Splicer redictor	http://spiceport.ebeb.und.edu/SpicingAnaryser.html	1998)
SpliceSiteFrame	http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm	-
SpliceView	http://zeus2.itb.cnr.it/~webgene/wwwspliceview.html	-
SROOGLE	http://sroogle.tau.ac.il/	(Schwartz, et al., 2009)

Table S1. Complete list of the bioinformatics resources used.

Name	Primer sequence	Restriction enzyme
Exon3_F	<u>ctcgag</u> cacatacaatgtggccttttca	XhoI
Exon3_R	tctagatattttgctgagcccattga	XbaI
Exon4_F	<u>ctcgag</u> ttttatggccactattcactgttt	XhoI
Exon4_R	ggatcccagaggcagtttacagaagatactca	BamHI
Exon5_F	<u>ctcgag</u> ctgcctagatgctgggaaat	XhoI
Exon5_R	ggatccttttaaaaagaagcaaggtctga	BamHI
Exon6_F	ccgctcgaggctgtggttcttgctttatgc	<mark>XhoI</mark>
Exon6_R	cgcggatcctctacaaaaatgaactgaatcaaca	<mark>BamHI</mark>
Exon11_F	ccg <u>ctcgag</u> cagagtgagcacttggcaac	XhoI
Exon11_R	cgc <u>ggatcc</u> ccattgaggacgtttgtctc	BamHI
Exon13_F	ccg <u>ctcgag</u> agcaaaatcacttcagcagttc	XhoI
Exon13_R	cgc <u>ggatcc</u> cacaaggcaatgatactgcaa	BamHI
Exon14_F	ccg <u>ctcgag</u> aagagcaacaaagetctgaca	<mark>XhoI</mark>
Exon14_R	cgcggatccctcttcccacagagggttca	<mark>BamHI</mark>
Exon16_F	ccgctcgagctgtgatcttggctttcttgtg	<mark>XhoI</mark>
Exon16_R	cgcggatccgctgcacatgctcacaattt	<mark>BamHI</mark>
Exon17_F	ccg <u>ctcgag</u> tgcagctcctgcagtttcta	XhoI
Exon17_R	cgcggatcctcaaatctcctgcccttttg	BamHI
Exon21_F	ccg <u>ctcgag</u> ccattaccaacaacacctcca	XhoI
Exon21_R	cgcggatccgtgaatcctatgacccagca	BamHI
Exon23_F	ccgctcgagtgcttctggcttgagcctat	XhoI
Exon23_R	cgcggatcccatttctcaacctggcgatt	BamHI

Table S2. Primer sequences used to generate the indicated minigene constructs.

Exon	ESEfir	nder3.0	FSJ	plice	Gen	eID	GenS	SCAN	HBOND	HS	SF	MaxEn	tScan	NetG	ene2	NNS	plice	Splice	ePort	SplicePr	edictor	SpliceSi	teFrame	Splice	View
number	Acc	Don	Acc	Don	Acc	Don	Acc	Don	Don	Acc	Don	Acc	Don	Acc	Don	Acc	Don	Acc	Don	Acc	Don	Acc	Don	Acc	Don
1	NA	9.89	NA	12.12	NA	444	NA	94	17.6	NA	96.51	NA	9.22	NA	97.9	NA	99	NA	143.98	NA	94.4	NA	89.76	NA	80
2	9.86	8.62	4.78	12.54	592	396	76	95	15.8	81.81	90.24	10.24	9.79	70.1	98.0	98	99	117.59	161.66	100	98.9	84.03	89.78	88	84
3	11.10	6.05	5.58	8.48	477	331	90	74	16.5	89.43	91.45	10.45	8.05	96.6	93.0	100	98	99.30	28.45	99.6	76.7	94.58	82.66	90	84
4	12.29	6.14	8.57	12.26	629	241	120	85	14.1	84.25	84.39	11.89	8.49	99.7	89.4	99	75	206.85	0.13	99.9	72.7	85.99	82.45	91	80
5	9.19	5.57	5.50	6.52	564	109	95	47	12.8	86.42	78.86	9.88	6.97	61.5	74.4	96	96	171.74	8.62	100	97.0	87.67	75.09	88	78
6	11.74	6.28	6.75	8.20	538	99	115	69	14.0	94.50	77.31	10.26	6.37	70.1	83.6	97	91	131.70	22.25	99.9	94.6	99.17	72.11	89	79
7	6.83	5.79	1.65	13.10	-6	236	36	98	14.0	85.59	83.72	3.78	9.49	15.9	87.9	0	97	-15.00	71.74	0	97.3	83.97	78.63	80	73
8	7.74	7.49	3.75	12.68	455	287	74	95	15.8	82.52	85.75	8.16	9.43	51.2	93.7	99	87	40.42	158.86	99.9	95.6	88.71	82.79	85	82
9	5.18	6.05	2.83	6.80	416	142	85	55	14.7	81.00	90.88	9.65	5.24	28.9	76.5	57	0	130.88	28.29	99.8	73.1	84.06	81.56	82	81
10	8.81	7.46	8.30	8.34	468	166	82	63	12.3	93.61	78.39	11.38	6.44	42.4	89.1	91	83	104.67	57.59	99.9	82.6	98.10	77.59	86	81
11	7.06	11.80	9.72	13.24	194	460	97	94	17.3	92.35	95.9	6.56	10.06	20.3	99.4	0	100	54.62	233.75	98.2	98.2	90.08	91.71	81	88
12	4.50	6.46	8.20	7.64	293	268	80	63	12.1	81.62	85.72	7.86	6.38	77.8	85.8	94	95	94.30	-4.74	99.4	92.4	82.58	82.95	83	85
13	8.88	8.13	9.30	12.54	535	396	89	95	15.8	86.86	90.24	8.89	9.79	89.7	96.3	97	99	-44.11	103.56	99.9	93.6	87.48	89.78	87	84
14	4.16	11.2	4.67	13.38	209	445	50	100	19.0	77.50	90.71	7.49	10.29	0.4	98.2	0	99	-27.49	170.28	97.2	99.8	80.86	89.63	79	87
15	6.18	9.58	6.10	13.24	188	460	50	94	18.6	92.27	95.9	6.71	10.06	5.0	99.5	83	100	-10.27	123.4	99.5	90.6	94.85	91.71	84	88
16	10.88	8.76	11.18	11.56	570	289	97	87	15.5	92.43	87.31	12.60	8.40	96.4	95.3	99	86	61.90	55.68	99.8	96.3	94.35	82.88	90	86
17	8.54	6.22	11.72	10.02	427	137	96	80	14.0	90.35	78.88	9.68	8.17	57.6	94.7	96	91	11.23	76.35	97.2	54.7	91.09	77.2	87	79
18	5.81	8.39	6.28	12.12	412	414	55	91	19.2	81.05	91.2	10.30	9.80	62.9	99.2	95	100	-9.91	137.63	99.8	98.4	84.67	90.02	82	85
19	7.08	7.86	9.53	11.42	359	306	81	86	13.8	87.59	87.91	10.02	9.11	44.6	96.9	88	97	87.91	33.54	99.6	83.1	87.63	84.51	87	82
20	1.49	3.93	0	10.02	16	47	0	79	14.0	85.54	76.54	4.35	7.64	0.1	93.9	0	91	39.18	-44.55	79.9	87.0	77.46	71.92	0	73
21	8.68	8.30	4.18	6.10	357	366	83	111	17.5	80.54	93.27	8.36	10.47	58.1	99.5	92	100	-10.43	192.75	98.4	96.5	82.40	87.4	84	85
22	10.41	8.15	9.50	8.62	456	269	75	76	13.6	92.37	90.7	9.83	6.65	81.0	95.2	98	0	185.12	-64.22	99.9	81.5	95.36	80.36	87	87
23	9.98	11.45	7.42	12.12	413	449	69	99	17.6	87.9	96.67	9.57	9.60	98.8	99.7	100	100	81.75	177.48	99.2	98.9	89.88	89.18	87	88
24	6.67	10.37	10.22	12.68	462	487	91	94	17.1	84.63	96.31	9.34	10.28	10.0	99.2	98	100	62.50	191.71	99.8	97.5	87.70	91.82	90	89
25	10.19	4.18	11.50	7.50	469	81	96	65	12.6	92.24	82.16	9.66	6.96	91.9	78.2	92	89	79.75	33.43	99.5	92.2	90.81	73.12	88	82
26	11.23	6.51	10.38	11.98	557	271	115	97	16.1	85.55	92.11	11.73	9.27	85.3	98.9	99	99	82.87	49.23	99.9	96.5	89.26	82.43	90	84
27	7.61	NA	4.60	NA	67	NA	50	NA	NA	84.1	NA	4.63	NA	81.2	NA	93	NA	80.77	NA	0	NA	74.15	NA	81	NA

Table S3. Score values for *CFTR* constitutive donor and acceptor splice sites calculated with thirteen different *in silico* tools. Values for acceptor and donor sites calculated with the GeneID, NetGene2, SplicePort and SplicePredictor algorithms were multiplied by 100 to facilitate the reading. The HBond algorithm calculates only the strengths of donor sites. For each strength analysis, the entire exonic sequence and part of its flanking intronic sequence (100 nt

up/down-stream) was used as input when possible; if not, the requested size of the constitutive splice site was used for the corresponding software. *NA*: not available. The complete list of references for each bioinformatics resource is available in Supp. Table S1.

	SROO	GLE			EX-SKIP	SKIP SKIPPY					
Exon/Intron number	BP	PPT	Exon size (bp)	ESS density	ESE density	ESS/ESE ratio (x100)	ESS density	ESE density	ESS/ESE ratio (x100)	Upstream ESS density (-100nt)	Downstream ESS density (+100nt)
1	NA	NA	185	30.81	79.46	38.78	NA	NA	NA	NA	NA
2	66.67	88	111	31.53	108.11	29.17	0.440	0.050	11.36	0.24	0.53
3	77.33	82	109	78.90	72.48	108.86	0.279	0.308	110.39	0.31	0.32
4	77.33	95	216	49.54	78.24	63.31	0.280	0.150	53.57	0.36	0.32
5	96.00	94	90	48.89	85.56	57.14	0.365	0.176	48.22	0.62	0.39
6	92.00	81	164	60.37	56.71	106.45	0.240	0.080	33.33	0.46	0.31
7	80.00	58	126	37.30	121.43	30.719	0.460	0.090	19.57	0.26	0.47
8	86.67	80	247	39.68	56.68	70	0.350	0.110	31.43	0.33	0.27
9	92.00	77	93	39.78	124.73	31.90	0.409	0.136	33.25	0.45	0.32
10	84.00	65	183	48.63	93.44	52.05	0.430	0.110	25.58	0.38	0.46
11	nd	Nd	192	51.04	88.54	57.65	0.380	0.160	42.11	0.26	0.26
12	84.00	78	95	53.68	116.84	45.96	0.522	0.178	34.1	0.2	0.41
13	77.33	93	87	71.26	90.80	78.48	0.293	0.293	100	0.36	0.26
14	80.00	75	724	33.98	103.04	32.98	0.330	0.110	33.33	0.47	0.31
15	77.33	70	129	59.69	72.09	82.80	0.360	0.250	69.44	0.51	0.56
16	77.33	78	38	78.95	52.63	150.00	0.364	0.152	41.76	0.27	0.46
17	82.67	87	251	43.82	67.73	64.71	0.360	0.120	33.33	0.31	0.28
18	84.00	72	80	46.25	56.25	82.22	0.253	0.160	63.24	0.40	0.50
19	100.00	71	151	66.23	74.83	88.50	0.380	0.180	47.37	0.30	0.31
20	nd	Nd	228	45.61	69.74	65.41	0.240	0.210	87.5	0.48	0.37
21	86.67	79	101	33.66	89.11	37.78	0.365	0.062	16.99	0.26	0.21
22	68.00	78	249	29.32	95.58	30.67	0.240	0.170	70.83	0.39	0.22
23	80.00	76	156	46.79	105.77	44.24	0.380	0.070	18.42	0.35	0.25
24	81.33	96	90	53.33	121.11	44.04	0.459	0.188	40.96	0.51	0.21
25	92.00	Nd	173	67.63	64.16	105.40	0.320	0.110	34.38	0.25	0.32
26	78.67	78	106	30.19	79.25	38.09	0.317	0.099	31.23	0.40	0.20
27	98.67	72	1754	59.12	78.73	75.09	NA	NA	NA	NA	NA

Table S4. Score values for putative *CFTR* **splicing signals calculated with SROOGLE, EX-SKIP and SKIPPY.** Values for the intronic polypyrimidine tracts and branch points were assigned to the corresponding downstream exons. Each exonic sequence with its entire flanking introns was used as input. For SROOGLE analyses, values calculated for the putative branch points were adjusted to 100%, which corresponds to the highest value calculated for *CFTR* introns with the Kol et al. algorithm (Kol, et al., 2005) to facilitate comparison. Values for the putative polypyrimidine tracts (PPT) were calculated with the Schwartz et al. algorithm (Schwartz, et al., 2009; Schwartz, et al., 2008) and adjusted to 100% as above. For EX-SKIP analyses (Raponi, et al., 2011), values for ESE and ESS densities corresponded to the total number of predicted ESSs or ESEs for each exon per 100 nucleotides. For SKIPPY analyses (Woolfe, et al., 2010), ESE and ESS densities were measured in the first and last 50bp of the exons, and when the exons were less than 100bp, densities were measured across its entire length; ESS densities were measured within 100 nucleotides upstream and downstream of each exon. Analysis of some splicing motifs could not be performed on exons 1 and 27 the first and last exons of *CFTR*, respectively. *NA*: not available. *nd*: not detected.

Exon number	Skill	SKOUOLE + EA-SKII +												
	Jim	ESEfinder3.0	FSplice	GeneID	GENSCAN	HBOND	HSF	MaxEntScan	NetGene2	NNSplice	SplicePort	SplicePredictor	SpliceSiteFinder	SpliceView
3	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
4	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Weak	Strong	Strong
5	Weak	Weak	Weak	Strong	Weak	Strong	Weak	Strong	Weak	Strong	Strong	Strong	Strong	Weak
6	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Weak	Weak	Strong	Strong	Strong	Weak	Strong
10	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
11	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Weak	Strong	Weak	Strong	Strong
13	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
14	Weak	Weak	Strong	Strong	Weak	Strong	Weak	Strong	Weak	Weak	Weak	Strong	Weak	Weak
15	Weak	Strong	Strong	Weak	Weak	Strong	Strong	Weak	Weak	Weak	Strong	Strong	Strong	Strong

SROOGLE + EX-SKIP +

Success rate of	fnrediction	10/13	8/13	0/13	10/13	8/13	0/13	7/13	10/13	8/13	0/13	7/13	8/13	10/13
23	Strong	Strong	Strong	Strong	Weak	Strong								
21	Strong	Strong	Weak	Strong	Strong	Strong	Weak	Strong						
17	Weak	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Strong	Strong	Strong	Weak	Strong	Strong
16	Strong	Weak												

Table S5. Comparing in vitro basal skipping and combined in silico predictions tools (core splicing signals + full-exonic ESE/ESS densities). Splicing skill

for each minigene construct was assigned as weak when the corresponding exon had a basal skipping rate superior or equal to 5% (Table 2). Score values for *CFTR* donor or acceptor splice sites were assigned as weak when found in the outliers under the lower inner fence (*CI90*) when compared with the median score of all *CFTR* exons (Figure S1). For combined *in silico* analysis, exons were considered weak if one of the indicated software tools tagged it as weak. A prediction was considered successful when the basal skipping skill was concordant to the strength of the splicing signals.

Exon number Skill														
Exon number	экш	ESEfinder3.0	FSplice	GeneID	GENSCAN	HBOND	HSF	MaxEntScan	NetGene2	NNSplice	SplicePort	SplicePredictor	SpliceSiteFinder	SpliceView
3	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
4	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Weak	Strong	Strong
5	Weak	Weak	Weak	Strong	Weak	Strong	Weak	Strong	Weak	Strong	Strong	Strong	Strong	Weak
6	Strong	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
10	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
11	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Weak	Strong	Weak	Strong	Strong
13	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
14	Weak	Weak	Strong	Strong	Weak	Strong	Weak	Strong	Weak	Weak	Weak	Strong	Weak	Weak
15	Weak	Strong	Strong	Weak	Weak	Strong	Strong	Weak	Weak	Weak	Strong	Strong	Strong	Strong
16	Strong	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak	Weak
17	Weak	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Strong	Strong	Strong	Weak	Strong	Strong
21	Strong	Strong	Weak	Strong	Strong	Strong	Weak	Strong	Strong	Strong	Strong	Strong	Strong	Strong
	Strong	Strong	Strong	Strong	Weak	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Strong
Success rate of p	rediction	9/13	7/13	9/13	9/13	7/13	9/13	7/13	10/13	6/13	8/13	6/13	8/13	9/13

 Table S6. Comparing *in vitro* basal skipping and combined *in silico* predictions tools (core splicing signals + surrounding ESE/ESS densities). Splicing

 skill for each minigene construct was assigned as weak when the corresponding exon had a basal skipping rate superior or equal to 5% (Table 2). Score values for

 CFTR donor or acceptor splice sites were assigned as weak when found in the outliers under the lower inner fence (*CI90*) when compared with the median score

 of all *CFTR* exons (Figure S1). For combined *in silico* analysis, exons were considered weak if one of the indicated software tools tagged it as weak. A prediction

 was considered successful when the basal skipping skill was concordant to the strength of the splicing signals.

SROOGLE + SKIPPY +

				EX-SKI	Р			SKIPPY		
		% of skipping	No. of ESS	No. of ESE	ESS/ESE	No. of ESE losses	No. of ESE gains	No. of ESS gains	LOR total	RC score
Exon	10									
WI		35%	89	171	0.52					
c.1240C>T	p.Gln414*	50%	93	160	0.58	4	0	3	4.118	1.899
c.1253A>G p	o.Asn418Ser	40%	91	173	0.53	0	2	0	-3.894	1.325
c.1270G>A p	o.Gly424Ser	69%	84	169	0.50	1	0	0	-0.568	0.824
c.1327G>T p	o.Asp443Tyr	32%	95	160	0.59	4	0	4	4.678	0.869
c.1331T>G	p.Ile444Ser	60%	84	169	0.50	0	0	0	-1.284	1.393
c.1355A>C p	o.Gln452Pro	4%	85	172	0.49	0	0	0	-1.284	2.001
c.1364C>A p	Ala455Glu	85%	94	178	0.53	0	4	1	-4.548	1.973
c.1366G>T p	o.Val456Phe	35%	94	170	0.55	1	0	2	2.006	1.794
Success rate of	f prediction		50%	50%	50%				12.5%	
Exon	13									
WI	-	15%	62	79	0.78					
c.1694A>G p	Asp565Gly	60%	64	66	0.97	4	1	0	-0.318	1.877
c.1727G>C p	o.Gly576Ala	100%	61	72	0.85	2	0	1	1.73	0.976
Success rate of	f prediction		50%	100%	100%				50%	

Table S7. Effects of point mutations on *in vitro* exon skipping and on ESE/ESS numbers and ratios or constitutions. EX-SKIP skipping was predicted when the number of ESS or ESS/ESE ratio increased or when the number of ESE decreased. SKIPPY skipping tendency was predicted when the LOR was significantly superior to 1. Rates of prediction success were calculated by confronting exon skills in minigenes and *in silico* prediction by each tool.

				EX-SKI	Р			SKIPPY		
		% of skipping	No. of ESS	No. of ESE	ESS/ESE	No of ESE losses	No of ESE gains	No of ESS gains	LOR total	RC score
Ex	on 3									
Ţ	WT	8%	86	79	1.09					
c.178G>T	p.Glu60*	32%	87	70	1.24	0	1	0	-2.962	4.597
c.202A>G	p.Lys68Glu	2%	86	84	1.02	0	0	0	-1.284	4.597
c.220C>T	p.Arg74Trp	39%	89	78	1.14	0	0	0	-1.284	7.662
c.223C>T	p.Arg75*	17%	89	79	1.13	0	1	2	-0.388	6.896
c.224G>T	p.Arg75Leu	37%	89	84	1.06	0	0	0	-1.284	7.662
c.224G>A	p.Arg75Gln	39%	87	79	1.10	0	0	0	-1.284	7.662
c.263T>G	p.Leu88*	3%	75	79	1.00	0	0	0	-1.284	0.764
c.263T>C	p.Leu88Ser	4%	79	79	0.95	0	0	0	-1.284	0.764
Success rate	e of prediction		100%	62.5%	87.5%				12.5%	
Ex	ion 4									
J	WT	0.7%	107	169	0.63					
c.328G>C	p.Asp110His	1.8%	107	163	0.66	3	0	0	0.749	1.414
c.350G>A	p.Arg117His	1.9%	107	170	0.63	0	1	0	-2.962	1.021
c.366T>A	p.Tyr122*	1.9%	105	170	0.62	0	0	0	-1.284	0.729
c.419C>T	p.Pro140Leu	1.3%	107	164	0.65	4	1	0	-0.318	1.615
c.443T>C	p.Ile148Thr	0.6%	104	167	0.62	1	1	0	-2.246	2.277
c.472A>C	p.Ser158Arg	0.3%	99	169	0.59	0	0	0	-1.284	1.759
Success rate	e of prediction		33%	50%	67%				50%	
Ex	con 5									
Ī	WТ	5 %	44	77	0.57					
c.496A>G	p.Lys166Gln	9 %	45	74	0.61	0	0	0	-1.284	2.268
c.509G>A	p.Arg170His	18 %	40	77	0.52	1	0	0	-0.568	2.308
c.533G>A	p.Gly178Glu	22 %	45	77	0.58	0	0	3	1.474	2.099
c.547C>A	p.Leu183Ile	10 %	48	77	0.62	0	0	0	-1.284	2.377
c.574G>A	p.Asp192Asn	8 %	46	75	0.61	1	0	5	2.75	1.672
Success rate	e of prediction		80%	40%	80%				20%	

Table S8. Effects of point mutations on *in vitro* exon skipping and on ESE/ESS numbers and ratios or constitutions. EX-SKIP skipping was predicted when the number of ESS or ESS/ESE ratio increased or when the number of ESE decreased. SKIPPY skipping tendency was predicted when the LOR was significantly superior to 1. Rates of prediction success were calculated by confronting exon skills in minigenes and *in silico* prediction by each tool.

	BEAS-2B	IB3-1	HT-29	HEK293	HeLa
WT	$7\%\pm3.2\%$	$6\%\pm0.9\%$	$12\%\pm0.6\%$	$13\%\pm2.8\%$	$9\% \pm 1.4\%$
c.220C>T	$39\%\pm0.3\%$	$19\%\pm4.0\%$	$32\% \pm 1.9\%$	$33\%\pm3.2\%$	$20\%\pm1.8\%$
c.223C>T	$17\% \pm 1.7\%$	$9\%\pm0.3\%$	$17\%\pm4.2\%$	$7\%\pm0.9\%$	$8\% \pm 1.2\%$
c.224G>T	37% ± 5.2%	$19\%\pm2.0\%$	$31\%\pm0.4\%$	$29\%\pm3.3\%$	$17\%\pm2.4\%$
c.224G>A	$39\%\pm0.3\%$	$18\%\pm1.0\%$	$34\% \pm 3.3\%$	$32\% \pm 1.1\%$	$25\%\pm4.7\%$

Table S9. Quantification of exon 3 splicing bearing the indicated mutation in different cell lines. The

percentage of exon 3 skipping, plus or minus SD, is represented. Experiments were repeated three to four times for each condition.

		HSF/ESEfind	ler3.0	ASSA	L
		Effect on SR protein RNA binding sites	Effect on hnRNPA1 binding sites	Effect on SR protein RNA binding sites	Effect on hnRNPH binding sites
Exc	on 10				
c.1240C>T	p.Gln414*	-2x 9G8		-	-
c.1253A>G	p.Asn418Ser	-1x Htra2-β +1x 9G8	-	-	-
c.1270G>A	p.Gly424Ser	+1x Htra2-β -1x 9G8	+1x (#)	-	-
c.1327G>T	p.Asp443Tyr	-1x 9G8	-	-	-
c.1331T>G	p.Ile444Ser	-	+1x (#)	-	-
c.1355A>C	p.Gln452Pro	+1x SC35	-	-SRp40 (ΔRi=-18.8)	-
c.1364C>A	p.Ala455Glu	+1x SF2/ASF (#)	+1x (#)	+SRp40 (ΔRi=6.3) +SF2/ASF (ΔRi=1.7)	-
c.1366G>T	p.Val456Phe	-	-	-SC35 (ΔRi=-6.8)	-
Exc	on 13				
		-1x SC35			
c.1694A>G	p.Asp565Gly	+1x SF2/ASF (#) +1x SRp40 (#)	-	-	-
c.1727G>C	p.Gly576Ala	+1x SRp55 (#)	-		-

Table S10. Effect of point mutations on putative RNA binding protein sites. Losses or gains of binding sites for four SR proteins (i.e. SF2/ASF. SC35. SRp40 and SRp55) and one hnRNP (hnRNPA1) are calculated with HSF and ESEfinder3.0. Changes in Relative information (Δ Ri) about the binding sites for three SR proteins (i.e. SF2/ASF, SC35 and SRp40) and one hnRNP (hnRNPH) are calculated with the Automated Splice Sites Analyses (ASSA) software and only those with a Δ Ri \geq 2 or \leq -2 are shown. # indicates when the implication of the predicted splicing factor was validated by *in vitro* modulation studies using a minigene approach (Pagani, et al., 2003a; Pagani, et al., 2003b).

		EX-SKI	P	ASS	SA
		Effect on SR protein RNA binding sites	Effect on hnRNPA1 binding sites	Effect on SR protein RNA binding sites	Effect on hnRNPH binding sites
Ex	con 3				
c.178G>T	p.Glu60*	-1x SC35 +1x 9G8	-	- SRp40 (ΔRi=-5.2) -1x SC35(ΔRi=-17.7)	+ hnRNPH (ΔRi=2.6)
c.202A>G	p.Lys68Glu	-	-	-	-
c.220C>T	p.Arg74Trp	-1x SF2/ASF (#) -1x SRp55	-	-SC35 (ΔRi=6.1)	-
c.223C>T	p.Arg75*	-1x SRp55	-	-	-
c.224G>T	p.Arg75Leu	-1x SF2/ASF -2x 9G8	-	-SC35 (ΔRi=-6.1)	-
c.224G>A	p.Arg75Gln	-1x SF2/ASF (#) -2x 9G8	-	-SC35 (ΔRi=-6.1)	-hnRNPH (ΔRi=-10.4)
c.263T>G	p.Leu88*	-	-	-	-
c.263T>C	p.Leu88Ser	+1 SRp55	-	+SRp40 (ΔRi=17.8)	-
<i>Ex</i>	con 4				
c.328G>C	p.Asp110His	-1x 9G8	-1x	-SC35 (ΔRi=-3)	-
c.350G>A	p.Arg117His	+1x Htra2-β	-	-	-
c.366T>A	p.Tyr122*	+1x SRp40	-	+SRp40 (ΔRi=6.3)	-
c.419C>T	p.Pro140Leu	-4x SF2/ASF	+1x	-	-
c.443T>C	p.Ile148Thr	+2x SRp40 +4x SF2/ASF	-	+SRp40 (ΔRi=5.4) +2x SF2/ASF (ΔRi=5.2)	
c.472A>C	p.Ser158Arg	-	-1x	-	-
Ex	con 5				
c.496A>G	p.Lys166Gln	-1x SRp40	-	-	-
c.509G>A	p.Arg170His	-1x SF2/ASF	-	-	-
c.533G>A	p.Gly178Glu	+1x 9G8	+1x	-	
c.547C>A	p.Leu183Ile	-1x SC35	-	-SC35 (ΔRi=-5.5)	-
c.574G>A	p.Asp192Asn	-	+1x	+ SRp40 (ΔRi=12.4)	

Table S11. Effect of point mutations on putative RNA binding protein sites. Losses or gains of binding sites for four SR proteins (i.e. SF2/ASF, SC35, SRp40 and SRp55) and one hnRNP(hnRNPA1) are calculated with HSF and ESEfinder3.0. Changes in Relative information (ΔRi) about the binding sites for three SR proteins (i.e. SF2/ASF, SC35 and SRp40) and one hnRNP (hnRNPH) are calculated with the Automated Splice Sites Analyses (ASSA) software and only those with a $\Delta Ri \ge 2$ or ≤ -2 are shown. # indicates when the implication of the predicted splicing factor was validated by *in vitro* modulation studies using a minigene approach.

SUPPLEMENTARY REFERENCES

- Blanco E, Parra G, Guigo R. 2007. Using geneid to identify genes. Curr Protoc Bioinformatics Chapter 4:Unit 4 3.
- Brendel V, Kleffe J. 1998. Prediction of locally optimal splice sites in plant pre-mRNA with applications to gene identification in Arabidopsis thaliana genomic DNA. Nucleic Acids Res 26(20):4748-57.
- Brunak S, Engelbrecht J, Knudsen S. 1991. Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol 220(1):49-65.
- Burge C, Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. J Mol Biol 268(1):78-94.
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31(13):3568-71.
- Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37(9):e67.
- Dogan RI, Getoor L, Wilbur WJ, Mount SM. 2007. SplicePort--an interactive splice-site analysis tool. Nucleic Acids Res 35(Web Server issue):W285-91.
- Kol G, Lev-Maor G, Ast G. 2005. Human-mouse comparative analysis reveals that branch-site plasticity contributes to splicing regulation. Hum Mol Genet 14(11):1559-68.
- Pagani F, Buratti E, Stuani C, Baralle FE. 2003a. Missense. nonsense. and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9. J Biol Chem 278(29):26580-8.
- Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, Baralle FE. 2003b. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum Mol Genet 12(10):1111-20.
- Raponi M, Kralovicova J, Copson E, Divina P, Eccles D, Johnson P, Baralle D, Vorechovsky I. 2011. Prediction of single-nucleotide substitutions that result in exon skipping: identification of a splicing silencer in BRCA1 exon 6. Hum Mutat 32(4):436-44.

- Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in Genie. J Comput Biol 4(3):311-23.
- Schwartz S, Hall E, Ast G. 2009. SROOGLE: webserver for integrative. user-friendly visualization of splicing signals. Nucleic Acids Res 37(Web Server issue):W189-92.
- Schwartz SH, Silva J, Burstein D, Pupko T, Eyras E, Ast G. 2008. Large-scale comparative analysis of splicing signals and their corresponding splicing factors in eukaryotes. Genome Res 18(1):88-103.
- Woolfe A, Mullikin JC, Elnitski L. 2010. Genomic features defining exonic variants that modulate splicing. Genome Biol 11(2):R20.
- Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11(2-3):377-94.