ELSEVIER

Mutation Research 411 (1998) 87–117

Reviews in Mutation Research

The association of nonsense codons with exon skipping

Carrie R. Valentine *

Division of Genetic and Reproductive Toxicology, National Center for Toxicological Research, HFT-120, 3900 NCTR Road, Jefferson, AR 72079-9502, USA

Received 3 November 1997; revised 3 March 1998; accepted 5 March 1998

Abstract

Some genes that contain premature nonsense codons express alternatively-spliced mRNA that has skipped the exon containing the nonsense codon. This paradoxical association of translation signals (nonsense codons) and RNA splicing has inspired numerous explanations. The first is based on the fact that premature nonsense codons often reduce mRNA abundance. The reduction in abundance of full-length mRNA then allows more efficient amplification during PCR of normal, minor, exon-deleted products. This mechanism has been demonstrated to explain an extensive correlation between nonsense codons and exon-skipping for the hamster *Hprt* gene. The second explanation is that the mutation producing an in-frame nonsense codon has an effect on exon definition. This has been demonstrated for the *Mup* and hamster *Hprt* gene by virtue of the fact that missense mutations at the same sites also are associated with the same exon-deleted mRNA. The third general explanation is that a hypothetical process takes place in the nucleus that recognizes nonsense codons, termed 'nuclear scanning', which then has an effect on mRNA splicing. Definitive evidence for nuclear scanning is lacking. My analysis of both nonsense and missense mutations associated with exon skipping in a large number of genes revealed that both types of mutations frequently introduce a T into a purine-rich DNA sequence and are often within 30 base pairs of the nearest exon boundary. This is intriguing given that purine-rich splicing enhancers are known to be inhibited by the introduction of a T. Almost all mutations associated with exon skipping occur in purine-rich or A/C-rich sequences, also characteristics of splicing enhancers. I conclude that most cases of exon skipping associated with premature termination codons may be adequately explained either by a structural effect on exon definition or by nonquantitative methods to measure mRNA, rather than an effect on a putative nuclear scanning mechanism.

Keywords: RT-PCR artifact; Premature termination codon; Stop codon; Reinitiation; mRNA abundance

1. Introduction

Attention was first drawn in 1993 to the association of premature nonsense codons in DNA with mRNA that lacks the exon containing the nonsense codon, referred to as exon skipping [1]. Since that time many reports have pointed out occasional associations between nonsense codons and exon skipping for a variety of genes (Table 1). Several explanations

Abbreviations: RT-PCR—reverse transcription-polymerase chain reaction; CHO—Chinese hamster ovary; Pu—purine; TCR—T-cell receptor

* Corresponding author. Tel.: +1-870-543-7329; fax: +1-870-543-7393; E-mail: cvalentine@nctr.fda.gov

1383-5742/98/$19.00
PII: S1383-5742(98)00010-6

Table 1
Characteristics of mutations that produce premature termination codons and are associated with exon deletions in mRNA that remove the nonsense codon (the first nonsense codon if a frameshift mutation). Mutations in splice-site consensus sequences are not included

Gene[a]	Total no. of exons	Exon deleted in cDNA[b]	Exon with nonsense[c]	Frame[d]	Base change (mutant or mutation no.)	Context (longest purine or A/C-rich sequence)[e]	Distance of mutation from end of exon[f]	mRNA abundance from mutant allele[g]	Proportion of mutant cDNA exon-deleted[h]	Missense nearby exon skipping[i]	Normal cells[j]	References
ADA	12	5	5	out	C → T (R142X)	AGGAGGGGG-AG-**C** GAGA	+63 (−54)	25–50%	half to all	−	yes	[2]
ANK	39	most of 38[k]	38	in	G → T (1669Glu → Ter)	AGAGAAT **G** AA	−19 from alternative splice site	normal	all	−	yes	[3]
CFTR	24	11 and 11,12	11	out	C → T (R553X)	GGAGGTCAA-**C** GAGCAAGAA	−23	low	almost all	−	−	[4]
		3[l] and 3,4	3	out, out	G → T (E60X)	agAGAATGGGAT-AGA **G** AGCTGGC-TTCAAAGAAAAA	+14	<15%	all	−	yes	[5–7]
		3 and 2,3	3	out, out	C → T (R75X)	GG **C** GATG (pur.) OR CAAAGAAAAA-TCCTAAACTC-ATTAATGCCC-TTCGG **C** GA (A/C)	+59 (−51)	low	all	−	−	
		1/2 of 19, all of 20, 1/2 of 21	20	out	G → A (W1282X)	GCAACAGTG-**G** AGGAAAG	−28	normal[m]	minor	−	−	[5–8]
DYS	79	part of 27	27	out	G → T	AAAGAAGAGG-CCCAACAAAAA-**G** AAGCGAAAG	+28	−	minor	−	−	[9]
		(3–7, genomic)	8	out	genomic frame-shift deletion exons 3-7 (patients 10,11)	−	−	25%-normal	(all, genomic)	−	yes	[10]
		3–9		in				−	minor	−	yes	
		2–7		in				−	minor	−	yes	
		2–9		out				−	least	−	yes	

E2	11	4 or 6 ($2–8^n$)	2	out, in	−AT (patient GM-612)	AAAC **AT** GTGGT-AA	+24, 25	normal[o]	50%	−	−	[11–13]
		4 or 6 ($2–8^n$)	6	out, in	G → T (patient GM-612)	AAAGTT **G** AAA	+115 (−103)	normal[o]	75%			
FAC VIII	26	19	19	in	G → T (patient Oxford 33)	GAGACAGTG **G**-AAA	+18	low	all	−	−	[14,15]
		22	22	in	C → T (patient Oxford 6)	AT **C** GAGGAAA	−27	low	half	−	−	
FACC	14	6	6	in	C → T (R185X)	A **C** GAG…(10 b) CCCACTTATTCCC	−33	−	half	−	−	[16,17]
FBNI	65	51	51	in	T → G (patient MS-7)	TA **T** GGAAGTG-GTGGGA	+26	25%	all	yes[p]	no	[1]
HL	9	2	2	in	G → T	G **G** AAG	−36 (+49)	−	28%	−	−	[18]
		2 and 2,3	2	in	G → T	G **G** AAG	−36 (+49)		some			[19]
Hpd	14	7	7	in	C → T (mouse strain III)	agAAAGCT **C** GA-GAACGGGGCG-CCAAAA	+7	normal	most	−	−	[20]
hprt Human	9											[21]
		2, 3	2	in	48–49delTT (LH-16C)	AGG (-**TT**) ATGA	+21, 22	−	50%	−	−	[22]
		2, 3	2	in	88insTGCTG (duplication of bases 84–88) (SP9-25)	A **TGCTG** AGGA-TTTGGAAAGGG	+57 (−46)	−	most[q]	−	−	[23]
		2, 3	2	in	88G → T (SP9-34)	T **G** AGGA	+61 (−47)	−	most[q]	−	−	[23]
		2–6	2	out	98delA (V-126-13)	GG **A** AAGGG	−37	−	all	−	−	[24]
		2, 3	3	in	139G → T (SP7-25)	**G** AACG	+5	−	most[q]	−	−	[23]
		2, 3	3	in	151C → T (253M78)	**C** GAGATGTGAT-GAAGGAGATG-GGGAGG[r]	+17	−	half	−	−	[25]

Table 1 (continued)

Gene[a]	Total no. of exons	Exon deleted in cDNA[b]	Exon with non-sense[c]	Frame[d]	Base change (mutant or mutation no.)	Context (longest purine or A/C-rich sequence)[e]	Distance of mutation from end of exon[f]	mRNA abundance from mutant allele[g]	Proportion of mutant cDNA exon-deleted[h]	Missense nearby exon skipping[i]	Normal cells[j]	References
Human												[21]
	2, 3		3	in	163A → T (253M92)	GAGATGTGATG **A**-GGAGATGGGAGG[r]	+29	–	10%	–	–	[25]
	2, 3		3	in	198T → A (LH-3C)	TG **T** G	+64	–	60%	–	–	[22]
	2, 3		3	in	207insG (V-133-11)	**AA** GGGGGG	+73	–	all	–	–	[24]
	2, 3		3	in	303delA (LH-72F)	AG **A** CTGAAGAG	−16	–	70%	–	–	[22]
	2, 3		3	in	307 A → T	G **A** AGAG	−12	–	all	–	–	[24]
	4		4	in	355G → T (mng113)	GGT **G** GAGATGA	−30	–	most[q]	–	–	[26]
	4–8		4	in	374T → G (HX92)	AACTT **T** AAC	−11	–	minor	–	–	[27]
	4–8		4	in	377delC (HX131)	AA **C** TGGAAAGg	−8	–	minor	–	–	[27]
	8		8	out	538G → T	T **G** GATTTGAAA	+6	–	most[q]	–	–	[28]
	8 and 2–6		8	out	538G → T	T **G** GATTTGAAA	+6	–	50% each	–	–	[22]
	8		8	out	544G → T (mng61; HX122; FTA2)	T **G** AAA	+12	–	most[q]	yes [29]	–	[26,27,30]
	8		8	out	550delC	GAAATT **C** CAGA-CAAG	+18	–	all	–	–	[24]
	8		8	out	589G → T (1smg22)	AAT **G** AATA	−21	–	20%	–	–	[31,26]
	8		8	out	594C → G	AATGAATA **C** TT-CAGGGA	+16	–	all	–	–	[24]
Hamster												[32]
	2 and 2,3 and 2–4		2	out, in, in	88G → T	**G** AGGATTTGG-AAAAGGTG	+61 (−47)	–	minor	yes	yes	[33]
	2 and others			out	97G → T	GAGGATTTG-**G** AAAAGGTG	−38	–	minor	yes	yes	[33]
					96 or 97delG	GAGGATTT **G** or **G** AAAAGGTG	−37/−38	–	minor	yes	yes	[33]
					101delA	GAGGATTTGG-AAAA (**-1A**) GGTG	−37/33	–	minor	yes	yes	[33]
	2 and 2,3 and 2–4		2	out, in, in	118G → T (240–5, NS2B)	AT **G** GAGTGA	−17	3%, 0%	15%	yes	yes	[34,36,37,33]
	2 and 2,3 and 2–4		3	out, in, in	139G → T (71–2)	**G** AAAGA	+5	2%	5%	–	yes	[36]

	3		151C → T (NS3A)	**C** GAGATGTCATG-AAAGAGAT GGGAGG[r]	+17	5%	minor	−	yes	[37]
	3		166G → T (NS3B)	GAGATGTCATG-AAA**G** AGATGG-GAGG[r]	+32	18%	minor	−	yes	[37]
	3		173delG (NS3C, D)	GAGATGTCATGA-AA-GAGAT (**-1G**)-GGGAGG[r]	+37/+39	7%	minor	−	yes	[37]
	3		205A → T (NS3E, F; 282-2B)	G **A** AGGGGGG-CTATAAA	+71	8%, 4%, 2%	17%	−	yes	[37,36]
	3		207insT (NS3G)	GAA (**+1T**) GGG-GGGCTATAAA	+72/73	5%	minor	−	yes	[37]
2–4	3	in	211delG (AA8QO-23)	AAGGGGGG (**-1G**)-CTATAAA	+73/78	−	some[s]	−	yes	[38]
2 and 2,3 and 2–4	3	out, in, in	217A → T (NS3H)	GAAGGGGGG-CTAT **A** AA	+83 (102)	7%	minor	−	yes	[37]
	3		233insT (NS3I, 238-1)	ACC (+1**T**) A	−86/87	1%[t]	39%	−	yes	[37,36]
	3		259A → T (NS3J, 282-5c)	AAAGCACTGAAT-**A** GAAATAGTGAT-AGA	−60	3%	18%	−	yes	[37,36]
	3		271A → T (NSK, L, M 282-9c)	AAAGCAGTGAAT-AGAAATAGTGA-T **A** GA	−48	2%	20%	−	yes	[37,36]
4 and 2–4	4	in, in	325C → T (HU/MNU17)	AATGAT **C** AGTCAAC-AGGGACATAAAAG	+7	−	most[q]	yes	no	[39]
	4		329C → G (251–7)	AGT **C** AACAGG-GGACATAAAAG	+11	30%	all	yes	no	[34,36]
	4		336delG (282–3C)	AACA (**-1G**) GGGG-ACATAAAAG	+16/+19	15%	all	yes	no	[34,36]
2 and 2,3 and 2–4	4	out, in, in	374T → A (NS4A, 282-8A)	TT **T** AACTGGAAAGg	−11	1%[t]	66%	−	yes	[37,36]
2 and 2,3 and 2–4	6	out, in, in	421A → T (NS6A, 282-8B)	GGT **A** AAACAATGC-AAA	+19	1%, 2%	4%	−	yes	[37,36]

Table 1 (continued)

Gene[a]	Total no. of exons	Exon deleted in cDNA[b]	Exon with non-sense[c]	Frame[d]	Base change (mutant or mutation no.)	Context (longest purine or A/C-rich sequence)[e]	Distance of mutation from end of exon[f]	mRNA abundance from mutant allele[g]	Proportion of mutant cDNA exon-deleted[h]	Missense nearby exon skipping[i]	Normal cells[j]	References
IDUA	14	2	2	in	C → A (patient D)	CCACACAGCCAGGA-TGACCAGTA **C** GTC-CTCA	+34	< 5%	major	–	yes	[40] [41]
IL-2Rγ	8	6	6	out	−1A (patient G)	AACCCACTCTGTGG-A **A**	+103 (−60) (exon 5)	reduced[u]	minor	–	–	[42] [43]
Mup	7	5	5	in	G → T (143Ter) G → T (172Ter)	CCGA **G** AACCAGA ACCTAT **CC** AATG-CCA	+4 −5	15% 90%	2% 2%	yes yes	yes yes	[44]
MNK	23	8	8	out	C → T	AAA **C** GAGAAATAA-GACAGTAAG	−14	–	some	–	–	[45], Genbank #X82342, L06133, X82356
NF1	60	37	37	in	C → A, C → G	ACACTTA **C** AACA	+36	–	half	–	–	[46], Genbank #L05367
OAT	11	6	6	out	G → A (W178X)	GGGAACTTCT **G** -GGGTAGGACG	+13	–	minor	–	–	[1]
		8	8	in	G → A (W275X)	AGAACTGGTAG-AT **G** GCTGG	+53 (−77)	–	minor	–	–	
PS-ALPHA	15	4	4	in	C → G (Ser62)	CACGTCAGT **C** AAC-TAA	+49 (−39)	low	minor	–	–	[47] [48]
Tg	42	9	9	in	C → T (cattle goitre)	GGAACT **C** GAAGT-GCGCTTGGAGAG	−28	low	half	–	yes	[49] [50]
WASP	12	3	3	in	C → T	GGGAA **C** AGGAG	+22	reduced	half	–	–	[51] [52]

Notes to Table 1:

[a] Human genes are given with all capitalized letters; genes from animals have the initial letter capitalized.

[b] Where 'and' is used, multiple species of cDNA with different exon deletions are found in the same cells. Full-sized species may also be present, see column 'Proportion of mutant cDNA exon-deleted'.

[c] Exon that contains the nonsense codon, whether created directly by the mutation or by a frameshift. If caused by a frameshift, then the first exon that contains a nonsense codon.

[d] 'in' means that after deletion of the exon in mRNA, the reading frame is maintained; 'out' means the reading frame is shifted.

[e] Mutated base in **bold**. Lower case letters indicate intronic sequences. Occasionally additional pyrimidine sequences are shown for a purine-rich sequence in order to show how the mutation changes purine content near a T nucleotide.

[f] Numbering is from the end of the exon deleted, '+' if from the 5′ end and '−' if from the 3′ end. Unless otherwise specified, the mutation is in the same exon as the nonsense codon.

[g] The steady state level of total mRNA (full-length and exon-deleted) as measured by Northern blot or amplified cDNA as determined by RT-PCR from the mutant allele relative to the unmutated allele, or if on the X chromosome, relative to cells with no mutation. This column does not indicate transcription from the gene. Multiple numbers refer to measurements by different methods or measurements on different mutants with the same mutation. 'Normal' indicates over 50%; 'low' indicates less than 25%.

[h] This is the steady state level of the exon-deleted transcript from the mutated allele compared to the total of full-length and exon-deleted transcript from the mutant allele. Again, this does not reflect the relative transcription, but steady-state levels of mRNA. If the exon-deleted transcript is more than 50% of the product from the allele containing the nonsense codon, it is designated 'major'; if less than 50%, minor. If quantitative information is available it is given. For autosomal genes, an exon-deleted species may be designated major even though it is a minor component relative to the normal transcript from the unmutated allele. For the X chromosome (e.g., *HPRT*) a high percentage of exon-deleted product may still be a minor product relative to the normal abundance in a parental line because the total steady-state transcript in the mutant line is reduced. Viewing the previous column with this column should make this distinction clear. An entry may say 'all' because a sensitive evaluation of the amount of normal species present was not done. Some data (hamster *Hprt*) are taken from column $\Delta 2{-}4/\text{Full} \times 100$ of Table 1 of Ref. [36]. This value is converted to reflect the proportion of the total, rather than the relative amount between species. The value in Table 1 of Ref. [36] ($\Delta 2{-}4 \div \text{Full} \times 100$) is the percent of exon-deleted product relative to full-length product. Therefore, the value given in that column is taken as the amount of exon-deleted product and 100 is taken as the amount of full-length product. The proportion of the total mRNA is then calculated by dividing the entry in Table 1 of Ref. [36] by that number plus 100. Numerically, the entry in Table 1 of this paper = % in Ref. [36] ÷ (% in Ref. [36] + 100%). For example, the $\Delta 2{-}4/\text{full}$ % entry [36] for mutant 238-1 is 64%; 64% ÷ (64% + 100%) = 39%, the entry in Table 1. This means that although the abundance of the exon-deleted species is two-thirds the abundance of the full-length species, it is only 39% of the total *Hprt* mRNA.

[i] 'yes' indicates that a missense mutation at or near the nonsense codon is associated with the same exon-deleted species of cDNA.

[j] 'yes' indicates that the same exon-deleted species of cDNA has been found in cells without mutation; 'no' is recorded only if a sensitive assay has been used to detect rare species.

[k] This alternatively-spliced form is normally a minor product (except in developing erythrocytes); in the mutant it is as abundant as the full-sized product in normal cells.

[l] The exon 3 deletion is found in both nasal epithelial cells and lymphocytes; the exon 3 and 4 deletion is found only in nasal epithelial.

[m] Two reports using allele-specific hybridization with oligonucleotides have indicated that the mRNA abundance in patients homozygous for this mutation is much reduced [53,5]. A third group has also found no mRNA using oligo-specific hybridization, but has found normal levels (twice the level of the normal allele) of CFTR mRNA in heterozygous W1282X carriers when detecting message by hybridization to PCR products covering exons 19–21 [54]. The W1282X mutation eliminates an *Mnl*I restriction site and the mutant allele can be specifically identified. The allele-specific oligo for W1282X may not hybridize efficiently.

[n] Since the exon 2–8 deletion removes mutations in both exon 2 and exon 6, it is not possible to tell from which allele the 2–8 deletion originated.

[o] Both alleles from this patient, one with a mutation in exon 2 and one in exon 6, are associated with the same exon deletions, either 4 or 6; the expression of mRNA from both alleles together is normal by Northern blot.

[p] The missense mutation [F + 2 (TAG)] results from a nonsense mutation whose frame has been changed by the insertion of two bp in a chimeric model construct [55]; the missense frame only partially eliminates the exon skipping.

[q] The mutation had to be detected from genomic DNA sequencing.

[r] The human sequence is a known exonic splicing enhancer [56]. The hamster sequence differs at two base positions, one of which decreases purine content.

[s] Full-sized cDNA is present, but the proportion of exon-deleted mRNA for this mutant is not given.

[t] For mutant 238-1, the relative abundance of the mutant full-length mRNA to the wild-type full-length species is 0.4% ([36], Table 1, column 'full'). The exon-deleted mRNA is 64% of the mutant full-length mRNA ([36], column $\Delta 2{-}4 \div \text{Full} \times 100$). Therefore, these two species together are about 1% (0.4% + 0.3%). For mutant 282-8A, the exon 2–4-deleted mRNA is twice the abundance of the mutant full-length species (also 0.4%) [36]. Therefore, these two species together are also about 1% (0.4% + 0.8%).

[u] Not quantitative.

have been offered (reviewed in Refs. [58,57,59]), including: (1) RT-PCR artifact, (2) sequence effects on exonic determinants of exon definition, and (3) nuclear scanning, in which a yet unknown nuclear entity recognizes codon sense in pre-mRNA or mRNA. When trying to identify a unifying mechanism, the association between nonsense mutations and exon skipping has been considered mysterious [58].

However, of these three different proposed explanations, two have been established for specific genes. An RT-PCR artifact based on the reduction in mRNA abundance caused by nonsense codons (Section 2.1)

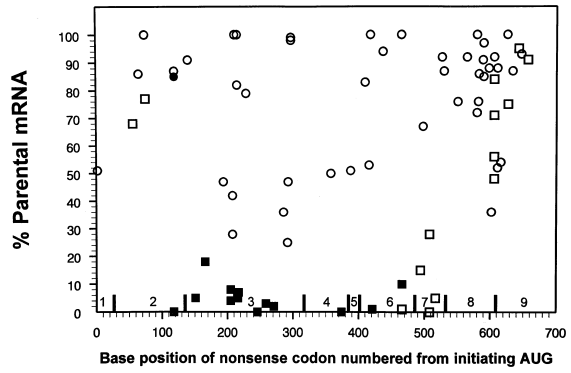

Fig. 1. Scattergram showing an association between nonsense codons and exon skipping that is caused by a nonsense-mediated reduction of mRNA abundance and an RT-PCR artifact. CHO *Hprt* nonsense mutants (squares) are compared with missense mutants (circles) for their relative mRNA concentrations, determined by Northern blotting, and for the presence of exon skipping (solid symbols), determined by preparative RT-PCR. Symbols for each mutant are displayed at the position in the cDNA of the nonsense codon generated by the mutation. If the mutation causes a frameshift, the position of the nonsense codon is downstream of the mutation. Exon-deleted mRNA species present in each mutant have deletions of exon 2, exons 2 and 3, and exons 2–4. The nonsense mutant mutated at base 466 has as a minor product only the deletion of exons 2–4 (the most abundant species). Vertical lines indicate exon boundaries and numbers indicate exon numbers. Except for one mutant, data are from Ref. [37]. The RT-PCR datum for the missense mutant at position 119 with exon skipping (G → T, mutant 237-1) is from Ref. [34] and the Northern blotting datum from Ref. [60]. PCR primers are 215 and 216 [61], which have their 3′ ends at −69 and 731, respectively. The actual abundance of exon-deleted products is not changed as a group between nonsense and missense mutants when measured by quantitative PCR. The apparent lack of exon-deleted products for missense and wild-type cells on agarose gels by screening PCR is the result of the abundant full-length mRNA suppressing amplification of the minor, exon-deleted products [36].

has been shown to be the cause [36] of an extensive association between premature termination codons and exon skipping in minor mRNA products for the *Hprt* gene of CHO cells (Fig. 1) [37]. The sequence change of mutations resulting in nonsense codons has also been shown to affect exon definition by virtue of the fact that missense mutations at the same sites are associated with the same exon deletions in mRNA for the *Mup* gene [44] and the human and hamster *Hprt* gene [36] (Section 3.2). A fairly lengthy list of missense mutations associated with exon-skipping is assembled in Table 2. Both nonsense and missense mutations that are associated with exon-skipping are often near the ends of an exon and frequently introduce a T nucleotide into a relatively purine-rich DNA sequence. The introduction of a U into a polypurine RNA sequence is also a characteristic of sequence changes that suppress exonic splicing enhancers [56,9].

I argue in this review that genuine associations between premature termination codons and exon skipping that removes these codons from mRNA are probably the result of the sequence change on exon definition (Section 3). Although this thesis is not proven, almost all of the current data is consistent with this interpretation, including some data that is used to support nuclear scanning models. A second explanation for the apparent associations between nonsense codons and exon skipping is an RT-PCR artifact (Section 2). Nuclear scanning is an active area of research, but there is still no direct evidence for its occurrence (Section 4). The primary purpose of this review is to guide the investigator faced with interpreting data that associates nonsense codons with exon skipping in reaching a plausible explanation.

2. Apparent association of nonsense codons with exon skipping based on RT-PCR artifact

2.1. Hprt hamster gene

Several investigators have suggested that the reduced abundance of mRNA accompanying nonsense codons could result in the detection of minor, exon-deleted products during RT-PCR [40,16,14,3,58]. A nonsense-mediated reduction of mRNA is a general occurrence that is found in all organisms examined, both prokaryotic and eukaryotic [58,57,59,74–76]. A

possible selective pressure for this effect of nonsense codons is to eliminate improperly spliced mRNA, which would contain nonsense codons if deleted or additional exons change the reading frame (Nagy and Maquat, 1998, Notes added in proof). Typically, the abundance of mRNA in internal exons is reduced to 0–30% of normal, depending on the gene [59,37,44,77–79] (squares in Fig. 1).

The most extensive correlation yet reported between nonsense codons and exon skipping for a single gene (hamster *Hprt*) (Fig. 1) [37] has been shown to be the result of an RT-PCR artifact, which is based on the difference in abundance between full-length and exon-deleted mRNA [36]. All of the minor, exon-deleted products found in nonsense mutants are present in wild-type cells at the same abundance when determined by quantitative RT-PCR. However, when using a preparative RT-PCR protocol, the abundant full-length mRNA species in wild-type cells suppresses the amplification of the minor, exon-deleted products 40-fold, which are then not readily recognized in agarose gels. For nonsense mutants, in which the abundance of full-length mRNA is usually reduced to less than 10% of normal (Fig. 1), more efficient amplification of minor products makes them visible on agarose.

Although the quantitative protocol used for the hamster *Hprt* study employed a restriction enzyme digestion to prevent the full-length mRNA from competing with minor, exon-deleted mRNA, simply diluting the product being amplified (about 80-fold, rather than the 1.5-fold incorrectly implied in Ref. [36]) is sufficient to achieve equivalent amplification of minor products from nonsense mutants and wild-type cells. This suggests a competition between abundant full-length and minor, exon-deleted mRNA. Competition for primers and reagents is a suggested mechanism for loss of minor products [55, Henegariu et al., 1997, Biotechniques 23: 504–511, added in proof]. The idea that minor species are not detected by PCR for normal cells has been questioned because exon-deleted species of simulated mRNA molecules can be accurately detected by RT-PCR at 0.01% the abundance of the larger species [80]. However, by quantitative RT-PCR of extracted cellular RNA, all of the exon-deleted species that appear in RT-PCR products of nonsense mutants of the hamster *Hprt* gene are present between 0.1% and 1% the abundance of full-length mRNA in mRNA from normal cells (Fig. 2) and only minor products exceeding 2% are recognized by the screening RT-PCR protocol [36]. The high concentration of the abundant mRNA (next paragraph), rather than the relative concentration, may be the determinant of suppression of minor species.

Another possible mechanism for the suppression of the amplification of minor products is rehybridization between exon-deleted and full-length species. Amplification of abundant species lags behind linear amplification at high concentrations of mRNA because of the C_0t effect, where a rehybridization reaction competes with amplification [81]. The common sequence between full-length and exon-deleted mRNA may contribute to a loss of exon-deleted products if abundant, full-length DNA rehybridizes with exon-deleted DNA. Fig. 2 shows that for the RNA input used in the screening studies of Fig. 1 (1 μg), amplification of the full-length species is beyond the linear range of amplification.

2.2. In-frame *vs.* out-of-frame exon-deleted products

Whether the exon deletion that removes the nonsense codon leaves the reading frame intact (in-frame) or changes the frame (out-of-frame) has a bearing on whether the appearance of exon-deleted products is explained by RT-PCR artifact. If the exon-deleted product is in-frame, the deletion removes the nonsense codon and the steady-state level of exon-deleted mRNA is presumably maintained as normal for that species. However, if the exon deletion is out-of-frame, then this species will also be subject to a nonsense-mediated reduction in abundance. If the exon-deleted species is already so minor that it is not detected by RT-PCR of mRNA from normal cells, then this further reduction in abundance may prevent it from being recognized even if the full-length species does not suppress its amplification. The 40-fold suppression of minor products by abundant full-length mRNA is comparable to the nonsense-mediated reduction of mRNA abundance (about 20-fold) [36].

We have observed the disappearance of a minor species in a hamster *Hprt* nonsense mutant for this reason [36]. When the abundance of the exon 2-deleted mRNA species drops 20-fold because of a position effect of a nonsense codon (mutant 71-2,

Table 2
Missense or silent mutations associated with exon deletions in mRNA

Gene and species	Exon deleted in cDNA	Exon with missense or silent mutation	Frame after exon deletion	Base change (mutant or mutation no.)	Context (longest purine or A/C-rich sequence) mutated base in bold	Distance of mutation from end of exon	Proportion of mutant cDNA exon-deleted	Reference
ADA	7	7	in	G → A (patient 2)	GGGGCC **G**	+37 or −36	50%	[62,63]
FBN1	51	51 (silent)	in	C → T	GGAAGTGGGATCAT **C**-GTGGGA	−26	−	[64]
GPIIIA	9	9 (silent)	in[a]	G → A	**G** GTGAGg	−6	all	[65]
HEXB	5′ half of 11 and all of 11	11	out, out	C → T	C **C** GGGCACAA	+8	14%, 7%	[66]
hprt Human								
	2,3 and 3–8	2	in	119G → T (7E5)	ATG **G** ACTAA	−16	all[b]	[67]
	2,3[c]	3	in	143 G → A (V-150-12)	GAAC **G**	+9	most	[24]
	4[c]	3	in	143 G → A (V151-7)	GAAC **G**	+9	all[b]	[24]
	2,3	6	in	446T → C (LH-56F)	**T** GG	−40	all[b]	[22]
	6	6	out	482C → A (mng109)	AAGGTCG **C** AAGg	−4	all[b]	[26]
	8	8	out	538G → A (mng113; mng141)	T **G** GATTTGAAA	+6	50%	[31]
	8	8	out	539G → T (19A9)	G **G** ATTTGAAA	+7	all[b]	[67]
	8	8	out	544G → A (15)	T **G** AAA	+12	5%, all[b]	[25,24]
	8	8	out	550–551 CC → TT	GAAATT **CC** AGACAAG	+18, 19	all[b]	[24]
	8	8	out	551C → T (34M30; 42M30) (EUG55)	**C** AGACAAG	+19	30% 50%	[25,67] [29]
	8	8	out	580G → T	CCCTT **G** ACTATAA	−37	some[d]	[68]
	8	8	out	589G → A (LH-18G;	AAT **G** AA LH-74G)	−21	40%	[22]
	8	8	out	590A → T (325-3-19)	AATG **A** A	−20	74%	[69]
	8	8 (silent)	out	597C → T	**C** AGGGA	−13	90%	[25]
	8	8	out	602A → T (sm97.1)	AGGG **A**	−8	some[d]	[67]

Hamster								
	2	2	out	34G → T (AA8QO-16)	AGT **G** ATGA	+7	all[b]	[38]
	2	2	out	83A → C	AATCACT **A**	+56 (−52)	minor	[33]
	2	2	out	119G → T (237-1)	ATG **G** AGTGA	−16	minor, 28-fold increase over normal	[36,33]
	4 and 2–4	4	in	329C → T HU/MNU18	AGT **C** AACAGGGGACA-TAAAAG	+11	all[b]	[39]
	4	4	in	329C → T (233-0-2UA)	AGT **C** AACAGGGGACA-TAAAAG	+11	all[b]	[36]
	4	4	in	332C → T	AGTCAA **C** AGGGGACA-TAAAAG	+14	all[b]	[70]
	4	4	in	338A → G	AGTCAACAGGGG **A** CA-TAAAAG	+20	all[b]	[70]
	3,4	4	out	355G → A HU/MNU3	GGT **G** GGGATGA	−30 (+37)	some[d]	[39]
	4	4	in	358G → T (SP65)	GGG **G** ATGA	−27	all[b]	[71]
MNK	20	20	in	G → A	AAGAGGAGGGGAAAC-GGGTAGCAATGGTGGG-AGAT **G** GAA	−102 (+103)	some[d]	[45], Genbank #X82353, L06133
Mup mouse	5	5	in	GA → TT	CCGA **GA** ACCAGA	+6, 7	minor	[44]
	5	5	in	A → T	CCGAG **A** ACCAGA	+7	minor	[44]
	5	5	in	G → C	CCGA **G** AACCAGA	+6	minor	[44]
	5	5	in	G → A	CCGA **G** AACCAGA	+6	minor	[44]
	5	5	in	TCC → GAG	ACCTA **TCC** AA	−8, 9, 10	minor	[44]
PBG	3	3	in	C → G (patient 3)	G **C** GTGGG	−22	most	[72], GenBank #M95623
ent MS-7)	GTGGGA							
PDHE1α	6	6 (silent)	in	A → G	GG **A** AAA	+45 or −49	50–80%	[73]

[a]An exon 9 deletion is an out-of-frame deletion; however, five bases of intron 9 immediately upstream from exon 10 are included because of the use of a cryptic splice site, making the over-all change in-frame.
[b]A single product contained an exon deletion.
[c]The same mutation is associated with different exon deletions in different clones.
[d]Proportion not given.

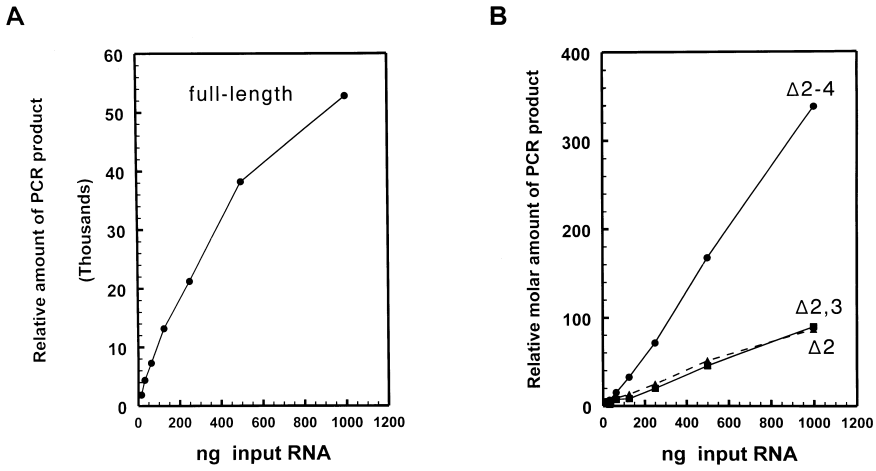

Fig. 2. Relative amounts of *Hprt* mRNA species from CHO cells detected by a quantitative RT-PCR protocol at different inputs of total RNA. The data is from Ref. [36] except that it has been corrected for the size of each species to reflect molar amounts between species rather than the bulk amount of DNA. (A) Full-length mRNA; (B) Exon-deleted mRNA species.

footnotes 3 and 5, Table 3, and Section 2.3), it is not visible on agarose gels by our standard screening methods even though the full-length species is reduced in this mutant to 5% of normal because of a nonsense mutation. Therefore, out-of-frame exon deletions are not expected to result in the appearance of minor, exon-deleted species in nonsense mutants if they were not visible in wild-type cells because the exon-deleted mRNA will also be reduced in abundance in the nonsense mutant, probably to a point below detection.

An important exception to this conclusion about out-of-frame exon deletions occurs when the first nonsense codon generated by the frameshift is close to the ends of the coding sequence. Here nonsense codons often do not have an effect on mRNA abundance (Section 2.3). Two of the three exon-deleted species found in nonsense mutants of CHO cells are

Notes to Table 3:

[a]The measured abundance below 10% [5] was determined by hybridization to allele-specific oligonucleotides; however, when using amplified cDNA as a probe, an abundance of CFTR mRNA is detected (200%) [54]. This implies that oligonucleotide hybridization may not be an accurate measure of mRNA abundance [6,54]. Since all the quantitative entries for the CFTR were measured with allele specific oligonucleotides, all of these entries may be subject to correction with further analysis. If the current values for the 3′ end of the sequence are valid, then it is unusual to have low mRNA for a nonsense mutation between two mutations with high abundance. The high abundance for a mutation almost 1000 bases from the normal termination is also unusual. There are five introns beyond this mutation.

[b]Relative to most abundant *Dhfr* mRNA of nonsense mutants

[c]Relative to full-length *Hprt* mRNA of wild-type cell. This mutation deletes the donor site of exon 2 and results in an exon 2 deletion in the major mRNA species with abundance greater than normal.

[d]Relative to exon 2,3-deleted *Hprt* mRNA of wild-type cells.

[e]Relative to abundance of exon 2-deleted mRNA in normal cells. This mutant has a nonsense mutation (nonsense for full-length mRNA) in exon 3. However, for the out-of-frame exon 2-deleted mRNA, this mutation produces a missense codon at what would be the first nonsense codon after the frameshift. The result is that the first premature nonsense codon of the exon 2-deleted mRNA with this mutation is 109 bases from the initiating AUG (normally the first nonsense codon is at 33 bases, first entry).

[f]Different cell isolates containing the same mutation.

[g]Relative mRNA half-life, not steady-state abundance

[h]A downstream, spliceable intron is required for the down-regulation of mRNA abundance. The distance to the intron is given rather than the distance to the normal termination codon.

[i]The distance that distinguishes those mutations associated with a reduction in mRNA abundance from those that are not is the distance from the last intron, intron 6. The required distance between the intron and the termination codon for reduced mRNA abundance is between 52 and 45 bases [91].

Table 3
Genes with greater mRNA abundance when a nonsense codon is near the 5′ or 3′ end of the coding sequence than when internal

Gene	5′ End of coding sequence		3′ end of coding sequence		Reference with mutant designation where appropriate
	Number of nucleotides from initiating AUG to nonsense codon	mRNA abundance relative to wild-type (% if number)	Number of nucleotides from normal termination to nonsense codon	mRNA abundance relative to wild-type (% if number)	
Bruton's tyrosine kinase gene	39	++			patient 54, [82,83]
(*BTK*) (human)	57	++			patient 50
	441	−			patients 55, 56
	526	−			patient 29
			463	−	patient 58
			421	±	patient 35
			273	+	patients 52, 53
cystic fibrosis transmembrane	117	< 30			Q39X [5,84]
conductance regulator gene	178	< 15			E60X
CFTR) (human)	223	< 10			R75X
	> 250	< 25			seven different mutations
			960	100	R1162X
			856	< 10	S1196X
			598	< 10[a], 200[a]	W1282X [5,54]
dihydrofolate reductase (*Dhfr*) (hamster)	~ 10	~ 100[b]			(figure 2 of Ref. [77]; GenBank
(*Dhfr*) (hamster)	~ 60	~ 100			#M13129, M19869) DU8
	90 exon 2 deletion	~ 80			
	~ 110	~ 10			
	~ 125	~ 10			
	> 125	~ 10			nine different mutations
			~ 130	~ 10	
			~ 130	~ 20	
			~ 100	100	
			~ 100	100	
dystrophin (*DMD*) (human)	132	> 25%			genomic deletion of exon 3′ [10,85]
	> 360	< 10%			8 different genomic deletions
fibrillin (*FBN1*) (human)			3473	6	5138ins 4 ([86,87]; GenBank #X63556)
			2672	25	6339T → G
			340	100	8236delGA
hypoxanthine phosphoribosyl transferase (*Hprt*, hamster)	33 (exon 2-deleted mRNA)	180[c]			264-0 [36,32]
	56	68			NS1A, figure 1 of Ref. [37]
	75	77			NS2A
	109 (exons 2,3-deleted mRNA)	30[d]			240–7 [36]
	109 (exon 2-deleted mRNA)	5[e]			71–2
	118	0			NS2B [37]
	151	5			NS3A

Table 3 (continued)

Gene	5′ End of coding sequence		3′ end of coding sequence		Reference with mutant designation where appropriate
	Number of nucleotides from initiating AUG to nonsense codon	mRNA abundance relative to wild-type (% if number)	Number of nucleotides from normal termination to nonsense codon	mRNA abundance relative to wild-type (% if number)	
	166	18			NS3B
	205	8, 4[f]			NS3E,F
	215	7, 7[f]			NS3C,D
			147	0, 0, 28[f]	NS7F, G, H
			138	5	NS7I
			48	56, 48, 71, 84[f]	NS8A, B, C, D
			30	75	NS9A
interleukin 2 receptor gene γ-chain (IL-2Rγ) (human)	241			moderate reduction	patient T [42,43]
	420		severe reduction	patient C	
		214		moderate reduction reduction	patient G
mouse urinary protein (*Mup*)	–		399	12	50Ter [44]
			120	15	143Ter
			33	90	172Ter
phytohemagglutinin of bean expressed in tobacco (*PHA*)			>350	<45	2 different mutations
			342	37[g]	([88], Fig. 6)
			174	160	
T-cell receptor-β			10 to intron[h]	strong down-regulation	construct A [89]
			8 to intron	down regulation	construct B
			2 to intron	weak down regulation	construct C
triose phosphate isomerase (*TPI*) (human)	6	86			1Ter [90]
	9	90			2Ter
	33	80			10Ter
	72	30			
			180[i]	15	189Ter [91]
			171	31	192Ter
			162	112	195Ter
			153	91	198Ter

in-frame deletions. The third is an out-of-frame deletion (exon 2 deletion) that is not subject to nonsense-mediated decay (footnote c, Table 3, mutant 264-0) [36]. This is presumably because the first nonsense codon produced by the frameshift is only 33 bases from the initiating AUG (footnote e, Table 3, and Section 2.3). A similar situation exists for an exon 2 deletion of the *Dhfr* gene ([77], mutant DU8). It is also possible that a frameshift mutation or exon deletion towards the end of a gene could shift the frame such that the first nonsense codon after the frameshift is past the 3′ boundary. This has been shown to be the case for exon 6 deletion of the human *TPI* gene, which restores normal abundance to mRNA when removing a nonsense codon [78]. A frameshift mutation in exon 5 of the IL-2Rγ gene may be also an example of this, since a single base deletion in exon 5 causes nonsense in exon 6, which is the exon deleted (Table 1). The first nonsense codon generated by the deletion is in exon 7, which

may be close enough to the end of the coding sequence (8 total exons) to escape the nonsense-mediated effect on abundance. Thus, this species could appear from a minor, normal product revealed when the full-length species is reduced in abundance.

Therefore, exon-deleted species that are expected to have increased detectability by RT-PCR when the full-length mRNA contains a nonsense mutation are either in-frame deletions or out-of-frame deletions that move the nonsense codon close to the ends of the coding sequence. There must also be a reduction of mRNA abundance for the full-length mRNA containing the nonsense codon. The generality of the down-regulation of mRNA by nonsense codons does caution that RT-PCR artifact must be considered as a possible cause of the appearance of minor exon-deleted PCR products when the exon deletion is in-frame or out-of-frame near the ends of the coding sequence. How close the nonsense codon must be to the ends of the coding sequence is addressed in Section 2.3.

2.3. Position effect of nonsense codons on mRNA abundance

The position of the nonsense codon relative to the ends of the coding sequence also has a bearing on whether the appearance of exon-deleted mRNA is explained by the RT-PCR artifact. The artifact that leads to an association between nonsense codons and exon-skipping is based on the fact that nonsense codons reduce mRNA abundance (Section 2.1). For several genes, the reduction in mRNA abundance occurs only when the nonsense codon is some distance away from the ends of the coding sequence (Fig. 1, Table 3). For the examples given in Table 3, the distance necessary from the 5′ end is about 100 nucleotides. For the 3′ end, the distance varies from approximately 100 to 400 bases. The spacing of documented mutants does not permit a clear boundary to be defined for most genes.

The gene for which the boundaries have been most clearly defined is the human *TPI* gene. The reason that premature nonsense codons upstream of this 5′ boundary do not cause reduced mRNA abundance appears to be because reinitiation occurs at a downstream AUG [90]. Because this AUG is in-frame, uninterrupted translation to the normal termination codon preserves normal mRNA abundance. Evidence for reinitiation has also been reported for frameshift deletions of the human dystrophin gene using exon-specific antibodies [92].

Since reinitiation is thought to be a relatively rare event [90], reinitiation may not explain the more general existence of a 5′ boundary for the nonsense-mediated effect on mRNA abundance indicated by Table 3. For example, the 5′ boundary of the hamster *Hprt* gene appears to be determined by the distance translated before premature termination rather than by reinitiation. Exon 2-deleted *Hprt* mRNA, which has an out-of-frame deletion, has its first nonsense codon 33 bases after initiation and is not subject to a nonsense-mediated reduction of mRNA abundance (footnote c, Table 3). Normal cells have a minor amount of exon 2-deleted mRNA and this amount does not decrease significantly in mutants with nonsense codons throughout exons 3–6 [36]. These mutants still have the nonsense mutation present in the exon 2-deleted mRNA species, and several mutations are still further than 100 bases from the 5′ end of the coding sequence even in exon 2-deleted mRNA. If reinitiation occurred at any of three in-frame AUGs in exon 3, the exon 4–6 mutations still present nonsense codons subsequent to this reinitiation that would be expected to result in reduced mRNA abundance. Reinitiation would have to skip the AUGs in exon 3 and occur in exon 6 in order to avoid a subsequent nonsense codon. The one exception to normal mRNA abundance for the exon 2-deleted species is a mutant (71-2) that has its first premature termination codon after 109 bases (footnote e, Table 3). This exon 2-deleted species has a 20-fold reduction in abundance relative to the same species in parental strains, typical of nonsense mutations. Therefore, increasing the distance translated before premature termination results in engagement of the nonsense-mediated downregulation of mRNA for the hamster *Hprt* gene.

The 3′ boundary for nonsense-mediated downregulation is in the penultimate exon of most genes. This was first noted for the *Dhfr* gene, in which none of the nonsense codons in the terminal exon or the 3′ end of the penultimate exon decreased mRNA abundance [77]. The 3′ boundary has been mapped accurately for two genes. For the human *TPI* gene it was found that the boundary is about 55 nucleotides upstream from the terminal *TPI* intron [93,91],

whereas, the T-cell receptor (TCR)-β gene possesses a boundary between only 2 and 8 bases from the terminal intron [89,59]. A mechanistic explanation for the difference in the 3′ boundaries that dictate nonsense-mediated downregulation is not clear.

The precise mechanism by which nonsense codons downregulate mRNA is also not known. Although for some genes this degradation may be attributed to a reduced half-life of cytoplasmic mRNA [58], many examples exist of genes whose mRNA associated with the nucleus is reduced to the same extent as cytoplasmic mRNA [59,94,95,44,96,89,97]. The reduction in nuclear mRNA abundance does appear to be the result of increased posttranscriptional degradation, not an altered rate of transcription [58,57,59]. The decreased nuclear abundance of mRNA for the human *TPI* and TCR-β gene is dependent on processes similar to translation [59,98–100]. Nuclear mRNA decay for the *TPI* gene may be coupled with the export of fully-spliced mRNA from the nucleus [78], and therefore a co-translational export model of mRNA decay has been proposed [58]. This model supersedes an earlier translational translocation model [95], for which predictions have proved incorrect [58,89]. In the co-translational export model, mRNA being exported from the nucleus begins translation at its 5′ end while still associated with the nucleus. Premature termination by ribosomes on the cytoplasmic side of the nuclear pore then initiates decay of the mRNA still associated with the nucleus. The co-translational export model has been further modified as a result of the finding that an intron must be downstream of a nonsense codon to trigger the full decrease in mRNA abundance [91,89,59]. Experiments have suggested that the intron itself is not the 'second signal' for nonsense-mediated decay [89] and therefore it has been proposed that the position of the intron in mature mRNA is 'marked' by covalent modification (e.g., methylation, or by retention of splicesomal proteins [91,89]).

Although models proposing that nonsense codons are recognized at the nuclear pore is attractive because nonsense codon recognition is achieved by cytoplasmic ribosomes, such models do not explain how nonsense codons can inhibit RNA splicing [101,102]. Therefore, alternative models in which transcripts are scanned in the nucleus proper by a hypothetical nuclear scanner have been proposed [58,57,89,59]. However, if inhibition of splicing is related to sequence effects on *cis* determinants of exon definition, then nuclear scanning mechanisms are not required to explain known effects of nonsense codons on exon skipping.

It has been suggested that the reversal of the nonsense-mediated reduction in mRNA abundance by protein synthesis inhibitors could be used generally to detect mRNA when screening nonsense mutants [103–105], which otherwise are often expressed at levels too low to be detected. This approach would also eliminate the basis for RT-PCR artifact affecting the amplification of minor mRNA species.

3. Structural effects of exonic mutations on exon definition

3.1. Exonic splicing enhancers

With the exception of the hamster *hprt* gene (Section 2.1), most examples of exon skipping associated with nonsense codons are for only isolated exons (Table 1). One explanation for this exon selectivity is that the sequence change of the mutations themselves are influencing exon definition [106] of weakly-defined exons independently of coding potential. Exonic regulatory elements have been defined that influence splice-site selection [107] and increase excision of introns at the 5′ end of the exon [108–112]. These sequences are generally purine-rich [108] or A/C-rich [110], but they do not have recognizable consensus sequences [112]. Purine-rich regulatory elements are bound by serine/arginine-rich (SR) proteins, which activate splicing of an adjacent intron [113,114]. Purine content alone does not dictate the efficacy of splicing enhancers since mixed purines can be more effective enhancers than either poly A or G and mutations that do not alter purine content can also affect enhancer function [112].

It has been shown that a U introduced into a polypurine RNA tract reduces enhancer activity [56]. An artificial polypurine sequence that mimics the exonic splicing enhancer of the dystrophin gene has reduced enhancer activity when a T is introduced into the DNA sequence [9]. When the T is introduced such that a nonsense codon is created, the suppression of enhancer activity is more pronounced that when the T mutation creates a missense codon.

However, the missense T is followed by two purines and one pyrimidine (GCG), whereas the nonsense T is followed by three purines (AAG). Therefore, the difference between the nonsense and missense mutation may result from differences in local purine content. The sequence TAGG has been shown to inhibit splicing of the exon containing it, with the dinucleotide AG having particular importance (this is the acceptor site consensus sequence, YAGG) [115]. In contrast, CAGG, TAAG, or TGAG do not inhibit splicing in the same context, suggesting that the T combined with the invariant AG dictates inhibited splicing in this context.

Because splicing enhancers are not readily identified from sequence information, it cannot be ruled out a priori that nonsense mutations are affecting an enhancer because the sequence context does not match that of a known enhancer. To establish a sequence as a known enhancer, it must be shown that the sequence provides a gain-of-function to an inefficiently-spliced, heterologous exon [112]. Very few sequences have been studied to this extent. To my knowledge, only two of the sequences shown in Table 1 is an exonic enhancer based on this criteria: the exon 3 sequence for the human *Hprt* gene (mutants 253M78 and 253M92) [56] and the exon 27 sequence for the dystrophin gene [9]. However, Table 3 lists many missense mutations that apparently do have an effect on exon definition. It is possible that certain exonic sequences are required for efficient splicing of their natural exons, but do not enhance splicing in other exons. I will refer to exon definition when the sequence has not been shown to act in a heterologous gene. The possibility of an effect on exon definition should be seriously considered for any nonsense mutation associated with exon skipping.

3.2. Missense mutations at the same or nearby site as nonsense mutations that are associated with exon skipping

In addition to the two examples of mutation in known exonic enhancers, there are several examples in Tables 1 and 2 for which missense mutations at the same or nearby nucleotides as the nonsense mutation also result in the appearance of the same exon-deleted species. These are exon 5 of the mouse *Mup* gene, exon 4 of the hamster *Hprt* gene, and exon 8 of the human *HPRT* gene. Although the exon-deleted species associated with missense and nonsense for the *Mup* gene are minor mRNA species, the hamster *Hprt* gene (329C → T or G) gives rise to a major mRNA species that has skipped exon 4 in response to either missense or nonsense mutations in exon 4. There are several examples of single base substitutions in exon 8 of the human *HPRT* gene where either nonsense or missense causes skipping of exon 8 in a large proportion of mRNA (538G → A or T; 544G → A or T; 589G → A or T) (Tables 1, 2 and 4). Thus, a single base substitution in an exonic sequence can result in loss of the exon in mRNA without premature termination. The exon skipping-inducing missense mutations in these genes do not affect known splicing regulators as defined by rigorous criteria (i.e., conferring regulation to a heterologous gene) but does indicate the existence of a *cis* acting determinant of exon definition.

3.3. Structural characteristics of nonsense mutations associated with exon skipping

Mutations producing in-frame nonsense codons occur within 30 bases of the nearest exon boundary for 40 of the 69 (58%) examples of exon skipping in Table 1. Six of the examples (11%) that are further away than 30 bases from the exon boundary have been shown to result from RT-PCR artifact, rather than true exon skipping (mutants NS3E, F, G, H, I, J, K of *Hprt* exon 3, Table 1, Section 2.1). After removing these six from the total, the percentage of nonsense mutations within 30 bases of the exon boundary associated with exon skipping is 63%. Of the 63 total examples then under further consideration, 45 are single base substitutions (at any location) and, of these, 33 (72%) introduce a T into a purine-rich sequence. Of the 12 base substitutions that do not introduce a T, seven increase the purine content following a T, two substitute the purine content following a T, and three increase the purine content in an A/C-rich sequence. All of these changes could affect splicing enhancers based on our knowledge of these elements (Section 3.1). Of 16 deletions or insertions, all but two occur in purine-rich sequences. Of those that do not, one (hamster *Hprt* mutant, NS3I) does not actually cause exon skipping since it results from an RT-PCR artifact, and the other (IL-2Rγ) is a good candidate for the same

Table 4
Characteristics of single base substitutions associated with exon skipping (Table 1,Table 2) compared to those not associated with exon skipping for exon 2 of the hamster [33] *Hprt* gene and exon 8 of the human [25,67] *HPRT* gene

With exon skipping			Without exon skipping		
Base change	Coding potential	Context longest purine-rich or A/C-rich sequence	Base change	Coding potential	Context longest purine-rich or A/C-rich sequence
Exon 2, Hamster					
83A → C	missense	AATCACT **A**	99A → G	missense	GAGGATTTGGAA **A** AGG
88G → T	nonsense	**G** AGGATTTGGAAAAGG	105G → T	missense	GAGGATTTGGAAAAGGT **G**
97G → T	nonsense	GAGGATTTG **G** AAAAGG	118G → C	missense	**G** GAGTGA
118G → T	nonsense	**G** GAGTGA	119G → C	missense	G **G** AGTGA
118G → T	nonsense	**G** GAGTGA	119G → A	missense	G **G** AGTGA
			130G → T	missense	G **G** ACAGgtaagtaaga
			130G → A	missense	G **G** ACAGgtaagtaaga
			131A → C	missense	GG **A** CAGgtaagtaaga
Exon 8, Human					
538G → T	nonsense	T **G**GATTTGAAA	536T → A	missense	G **T** TGG
538G → A	missense	T **G**GATTTGAAA	539G → A	missense	G **G** A
539G → T	missense	G **G** A	543T → G	missense	**T** GAAA
544G → A	missense	**G** AAA	551C → T	missense	C **C** AGACAAG
544G → T	nonsense	**G** AAA	551C → G	missense	C **C** AGACAAG
550–551CC → TT	missense	GAAATT **CC** AGACAAG			
551C → T	missense	C **C** AGACAAG	554A → T	missense	CCAG **A** CAAG
580G → T	missense	CCCTT **G** ACTATAA	563T → A	missense	G **T** TGTAGGA
589G → T	nonsense	AAT **G** AATA	568G → C	missense	A **G** GA
594C → G	nonsense	AATGAATA **C** TTCAGGGA	574G → C	missense	AGGATAT **G**
597C → T	silent	**C** AGGGA	580G → C	missense	**G** ACTATAA
602A → T	missense	AGGG **A**	585T → G	nonsense	A **T** AA
			599G → A	missense	A **G** GGA
			601G → T	missense	AGG **G** A
			606G → C	missense	**G** AA

Exon 2 extends from base 28 to 134; exon 8 is from base 532 to 609. Lower case letters are intronic sequences. Mutations in the −1, −2, or −3 position of the donor site are not included. Amounts of exon-deleted product vary from minor to predominant. One mutation (551C → T) appears in both columns since the same investigators found two different results with two different human donors.

RT-PCR artifact (Section 2.2). Deletions and insertions might also be expected to have a structural effect on exon definition based on the different length of the exon.

3.4. Structural characteristics of missense mutations associated with exon skipping

There is a sizable number of missense (or silent) mutations that are associated with exon skipping (Table 2). Many of these missense mutations result in accumulation of an abundant mRNA that has lost the missense-containing exon, again indicating that a single exonic base change can have a dramatic effect on exon definition. Many examples are from the *Hprt* gene, which may reflect the fact that this gene has been well studied [116] and even has a computerized database for human mutations that identifies exon skipping [117].

The characteristics noted in Section 3.3 for nonsense mutations associated with exon skipping are similar for missense mutations. Almost all missense mutations that increase the accumulation of exon-skipped mRNA are within 30 bases of the nearest exon boundary (30 of 36 total, 83%) compared to 63% of nonsense mutations (Section 3.3). Although this indicates that *cis* determinants of exon definition are usually near the ends of exons there are three examples of missense mutations in the middle of

exons that are associated with exon skipping. One of these is associated with a large proportion of deleted mRNA (*PDHE1α*). Therefore, it cannot be ruled out that mutations within the central region of an exon have an effect on exon definition.

Fifty percent (18 of 36) of the missense mutations associated with exon skipping introduce a T, and all but one of these substitutions into a purine-rich sequence. The one exception is the substitution of a T in an A/C-rich sequence. Of the remaining 18 mutations not introducing a T, two increase the purine content of a purine-rich sequence, 10 substitute one purine for another within a purine-rich sequence, and four substitute nucleotides in A/C-rich sequences. Although missense mutations associated with exon skipping are more localized toward the ends of exons than nonsense mutations, missense mutations show somewhat greater diversity in sequence changes, probably because they are not constrained by the nonsense codon consensus sequence, TPuPu.

The fact that the majority of missense mutations that cause exon skipping are at the ends of exons implies that exon termini harbor splicing-regulatory elements. If this is true, then if nonsense mutations act in the same manner as missense mutations, one would expect a similar skewed distribution toward exon termini. The higher frequency of nonsense mutations further than 30 bases from exon termini associated with exon skipping suggests that other mechanisms besides an effect on splicing-regulatory elements is conferred by nonsense codons. One possibility is a nuclear scanning mechanism (Section 4) and the other is an RT-PCR artifact (Section 2).

Two reported missense mutations in the human *HPRT* gene are unusual for being associated with the deletion of distant exons. A mutation in exon 3 is associated with the deletion of exon 4 in one isolate (143G → A, V151-7) and a mutation in exon 6 is associated with the deletion of exons 2 and 3 (446T → C) (Table 2). The exon 6 mutation occurs in neither a purine-rich nor A/C-rich sequence. These apparent examples of *cis* elements acting at a distance are intriguing, but it remains to be determined if other closer mutations are responsible for the effect.

The predominance of mutations within purine-rich sequences could reflect simply that mutations are more common in these sequences. The database of mutations inactivating the human *HPRT* gene shows that the most frequently isolated mutations occur at G bases on the non-transcribed strand (mRNA), frequently within a GG dinucleotide ([117], version 6). Conversions from G to A or T are much more frequent than to C. Mutations known to be associated with exon skipping by RT-PCR artifact (Section 3.3) also introduce a T into a purine-rich sequence. To address this issue, Table 4 tabulates mutations in exons 2 and 8 of the *Hprt* gene that *are* or *are not* associated with exon skipping, compared within the same study. Of 17 mutations associated with exon skipping, 14 introduce a T into a sequence of purines (82%). Of 22 mutations not associated with exon skipping, five introduce a T into a purine-rich sequence (23%). Both sets of mutations occur in purine-rich sequences. This suggests that the introduction of a T into a purine-rich sequence may be more frequent among exonic mutations that affect exon definition than those that do not. This is consistent with known effects on exonic enhancers (Section 3.1).

Stem-loop structures at the exon/intron boundaries have been proposed to distinguish between mutations that are associated with exon skipping and those that are not [33,25]. For example, hamster *Hprt* exon 2 mutations that stabilize a proposed stem structure are associated with exon skipping [33]. In contrast, mutations that disrupt the proposed base pairing in human *HPRT* exon 8 are associated with exon skipping [25]. However, further mutations identified in exon 8 associated with exon skipping have not disrupted the stem [118,24]. The choice of alternative acceptor sites has been shown to depend upon single exonic bases that are proposed to act through stem-loop structures [35,119]. It is possible that secondary structures of this type are part of the recognition of exonic splicing enhancers or general exon definition. The fact that enhancers are purine- or A/C-rich but have no consensus sequence suggests that their functionality resides in their possessing an appropriate secondary structure.

At first glance, it appears that nonsense codons cause exon skipping more frequently than do missense codons (since there are more entries in Table 1 than Table 2). There are several possible explanations for this. First, since considerable attention has

been placed on the association of nonsense codons with exon skipping, nonsense codons are more likely to appear in the title and abstract of an article and hence are more readily retrievable from the scientific literature. Second, since nonsense codons themselves are purine-rich, having the consensus sequence TPuPu, the introduction of a T in an enhancer sequence (purine-rich) may produce more nonsense codons than in other sequence contexts. Third, purine-rich sequences are more likely targets for mutagenesis (as mentioned earlier). Lastly, mutations resulting in nonsense codons may be over-represented because they have a dramatic impact on protein function. A study of point mutations selected in cell culture in the hamster *Hprt* gene targeted mostly purines and included 35 nonsense mutations and 48 missense mutations [37], indicating that nonsense mutations are frequently isolated when lack of protein function is the selective criteria.

The striking conclusion of this analysis is that almost all mutations associated with exon skipping affect sequence tracts that are likely to be splicing enhancers, by virtue of their being purine- or A/C rich. Most exon-skipping-inducing mutations fall into the following categories: (1) they introduce a T into a purine-rich sequence, (2) they increase or substitute the purine content following a T, (3) they increase or substitute the purine content in a purine-rich sequence, and (4) they occur in an A/C-rich sequence. The sequence contexts of missense mutations associated with exon skipping are more varied than those of nonsense mutations, suggesting that nonsense mutations are well within the sequence changes known to affect exon definition. Although this analysis does not prove that nonsense mutations act by altering the function of exonic enhancers, it does show that these mutations are consistent with the known effects of single base changes on enhancers (Section 3.1) and of missense mutations on exon definition (Section 3.2). Thus, it is a reasonable hypothesis for investigation that the association of nonsense mutations with exon skipping is frequently the result of sequence effects on exon definition. The location of nonsense codons (as well as missense codons) that are associated with exon skipping may be a rich source of candidate sequences for the identification of exonic splicing enhancers or *cis* determinants of exon definition.

4. Nuclear scanning

Nuclear scanning refers to an hypothesized ability of an unknown nuclear entity to determine the coding potential of nuclear RNA. This process was first proposed to explain the fact that nonsense mutations in the *Dhfr* gene were associated with reduced mRNA levels in the nucleus unless the nonsense codon was in the final exon or at the end of the penultimate exon [95]. Nuclear scanning has subsequently been invoked to account for exon skipping in fibrillin gene mRNA associated with nonsense mutations at a single site in exon 51 [55]. Nuclear scanning proposes the intriguing idea that the recognition of a nonsense codon in the nucleus can regulate splicing of an upstream intron. Various nuclear scanning models have been proposed that describe scanning before, during, and after splicing [58,57,89,59], but none has been established as a known mechanism. Nuclear scanning models will not be reviewed here. Instead, I will point out that some data often cited in support of nuclear scanning models is also consistent with an effect of the mutation on exon definition.

4.1. Fibrillin gene

The strongest evidence for nuclear scanning is a site-specific mutagenesis study of fibrillin gene exons in a transfected mini-gene construct [55]. It is sometimes assumed in the scientific literature that these experiments have established the existence of a nuclear scanning mechanism [120,97,9]. However, a silent mutation in the same exon as the nonsense mutations also results in exon skipping of that exon [64]. The possibility should be considered that the nonsense mutations are also acting directly on exon definition.

The first nonsense mutation associated with exon skipping in the fibrillin gene created a TAGG sequence by introducing a G at position +26 of exon 51 [1]. This sequence has been shown to be an inhibitor of exonic splicing enhancers [115] (Section 3.1). In cell culture, all three nonsense codons at the same site in a transfected mini-gene construct result in aberrant splicing, including one product containing a portion of intron that introduces a nonsense codon [55]. One missense mutation did not result in aberrant splicing (TAC); the only missense mutation

with the nonsense consensus sequence TPuPu, TGG, was not tested. These results could, therefore, be explained by a structural interpretation that exon skipping is related to forming a sequence with a T followed by purines (Sections 3.3 and 3.4).

The experiments that have been so convincing in support of nuclear scanning is the fact that a 2-bp frameshift in the mini-gene construct, several hundred bases distant in the adjacent exons, which converts the nonsense sequences to sense, eliminates aberrant splicing for two of the nonsense codons, TGA and TAA [55]. For the third, TAG, which forms a cryptic acceptor site ([121] and above), the frameshift only partially restores normal splicing. This is interpreted as a structural effect.

The fact that a silent mutation 26 bases from the other end of the same exon (3′ end of exon 51) [64] and the TAG sequence both cause skipping of exon 51 suggests that this exon is weakly defined. It seems possible that long range structural interactions caused by the 2-bp insertion and deletion could affect definition of the exons in which they occur, and consequently of the adjacent exon, particularly since these results are obtained for an in vitro mini-gene construct. The 2-bp insertion is 53 bp from the 3′ end of exon 50 and the 2-bp deletion is 89 bp from the 5′ end of exon 52. Intron 50 is 380 bases and intron 51 is 244 bases. Although reasonably well-removed from the ends of their exons, it is possible the sequence changes as well as the change in the length of each exon affects exon definition across the relatively short introns. It is known that a polymorphism at the +18 position of exon 9 of the platelet fibrinogen receptor gene, which does not cause exon skipping, is required for a mutation at the −6 position of exon 9 to result in skipping of exon 9 [65]. These two sequence changes also result in the use of a cryptic acceptor site in exon 10. Also, an exonic, single base deletion adjacent to the acceptor site of exon 5 of the *DHPR* gene results in deletion of exon 4 [122]. Thus, exon definition can be affected by alterations at the other end of the same exon or in an adjacent exon. Although this type of interaction is not characterized, it is no more difficult to imagine than nuclear scanning.

With respect to the entire fibrillin gene, all of the mutations causing exon skipping are in intronic sequences except for the three nonsense mutations at position +26 of exon 51 and the one silent mutation at position −26 of exon 51 [86,64]. Of five locations where frameshift or nonsense mutations are shown to be associated with reduced abundance of mRNA in the fibrillin gene, only the site in exon 51 has been reported to cause exon skipping [86]. It has been argued that exon skipping may not have been examined for other mutations of the fibrillin gene [121]. However, it has been pointed out that exon skipping is one of three major categories of *FBN1* mutations [123], constituting 10% of mutations (60 total) when screened by long RT-PCR [64]. In this screening study, the only exonic mutation that caused exon skipping was the silent mutation in exon 51. These facts argue that the effects of the +26 nonsense mutations are site-specific, rather than nonsense-specific. It has also been argued that not all nonsense mutations produce an effect on mRNA abundance and, therefore, not all nonsense codons should be expected to have an effect on exon skipping [121]. However, the effect of nonsense mutation on mRNA abundance is consistent for several genes within outside boundaries (Section 2.2, Table 3) [124,77,37]; this has not been demonstrated for the correlation between nonsense mutations and exon skipping (except for hamster *Hprt*, which results from RT-PCR artifact, Section 2.1).

The nuclear scanning model proposed for the fibrillin gene requires either that the reading frame be recognized in pre-mRNA across introns or that each exon is scanned locally for any open reading frame. This latter variant of nuclear scanning predicts that frameshift mutations that produce the first nonsense codon in the next exon would cause no exon skipping because both exons would have an open reading frame. The IL-2Rγ receptor gene (Table 1) does not fit this prediction since a single base deletion in exon 5 causes nonsense in exon 6, which is the exon deleted (Section 2.1).

The exclusion from mRNA of an exon containing a nonsense codon could confer a selective advantage if the deletion maintains the reading frame. A milder clinical condition resulting from in-frame exon skipping has been observed both for the human dystrophin gene [10] (Section 5.3) and for adenosine deaminase deficiency [62] (patient 2) than if protein activity is completely abolished by an out-of-frame deletion or by a particularly debilitating missense

mutation. Therefore, a nuclear mechanism that scans pre-mRNA for the purpose of excluding exons with nonsense codons from mRNA may have a physiological purpose separate from destroying non-functional mRNA in order to save energy or avoid toxic protein products [59]. The use of the cryptic donor site within intron 50 (57 bases into intron 50) for all three nonsense codons in the fibrillin mini-gene construct introduces a second nonsense codon [55]. This indicates that removal of nonsense is not the driving force behind altered splicing. If nuclear scanning affects only poorly defined exons, and then simply causes ambiguity in splice-site selection, there seems to be little selective advantage if the nonsense codon is not removed. Because of the limited appearance of exon skipping, nuclear scanning as a cause of exon skipping makes more sense if exon skipping is a by-product of scanning for the purpose of reducing mRNA abundance, rather than the ultimate purpose for scanning. Thus, the sporadic incidence of exon skipping would accompany the more regular occurrence of reduced mRNA abundance, if nuclear scanning is an actual process.

4.2. Minute virus of mice

A study of nonsense-causing mutations of the parvovirus, MVM, is also frequently cited in support of nuclear scanning models [102]. Various nonsense codons are associated with the inhibition of splicing of a large intron mRNA precursor (not exon skipping) when they are in an active reading frame, but not when they are out of the active frame. However, this study includes no missense mutations and many of the nonsense codons are generated by insertions of up to 16 base pairs. The results could be interpreted as structural effects of the mutations related to their proximity to splice sites, rather than the activity of nonsense codons. It is now known that sequences within a small intron must be present initially, in the proper orientation, in order for the large intron (Fig. 2A) to be spliced efficiently [125]. There is evidence that the size of the exon between these two introns has an effect on splicing of the large intron. All of the double mutants in the splicing study [102] are between these two introns. It seems possible that insertions and deletions in this region might also have an effect on splicing.

5. Comments on selected genes

Some genes that are well studied, or have unusual properties, merit further comment. In this section, data from individual genes relating nonsense codons and exon skipping are discussed.

5.1. Human HPRT gene

The human *HPRT* gene is highly studied as a molecular target for mutagenicity assays because mutant lymphocytes from blood specimens can be isolated in cell culture with selective media [116]. Exon-deleted species have been observed in human cells with nonsense mutations [118,24] (Table 1) and, in several of these cases, the major mRNA species is exon-deleted. Missense mutations in exons 2, 3, 6, and 8, are reported to be associated with the deletion of single or multiple exons in mRNA. These examples are likely to reflect sequence effects on exon definition.

The deletion of exons 2–6 (out-of-frame) occurs for nonsense mutations primarily when the mutation causes an out-of-frame structural event, such as a 67 bp inclusion of intron 5 from a donor site mutation (personal communication, J.P. O'Neill). This characteristic suggests that the appearance of the exon 2–6-deleted species may be the result of structural effects of mutations on splicing. Although the exon 2–6-deletion is an out-of-frame deletion, the size of the deleted segment may prevent the frameshift from reducing mRNA abundance. The first stop codon for the human exon 2–6 species occurs after 66 bases (positions 522–524 of cDNA sequence numbered from AUG). Therefore, translation may terminate before an effect on messenger abundance takes place (Section 2.3). In addition, this stop codon is at the end of exon 7, which may be close enough to the end of the coding sequence to escape an effect on mRNA abundance. Since this species may, then, have normal stability, it could also appear in nonsense mutants by RT-PCR artifact if it is a minor species in normal cells. Either the exon 2–6-deleted species is more abundant in normal human cells than hamster, or mutations in the human gene are more likely to result in omission of these exons during splicing, since the appearance of exon 2–6-deleted mRNA is more frequent in human mutants than hamster. For hamster, exon 2–6-deleted mRNA is seen only after

two rounds of PCR [36], whereas one round is sufficient mRNA to visualize other RT-PCR products for hamster.

Approximately 10% of human T-cell mutants with nonsense mutations in exons 1–4 and both T-cell and lymphoblasts from Lesch–Nyan patients contain minor species of *HPRT* mRNA with the deletion of exons 2 and 3 (in-frame) (personal communication, J.P. O'Neill). The large proportion of mutants with exon 2,3-deleted mRNA suggests that this is a normal mRNA species relatively increased during RT-PCR because of a nonsense-mediated effect on full-length mRNA abundance (Section 2.2). Although the most abundant exon-deleted species for the hamster *Hprt* gene is the exon 2–4-deleted mRNA (Fig. 2), the exon 2,3-deleted mRNA may be the most abundant for the human gene. However, it has not been demonstrated at this time that exon-deleted species of the human *HPRT* gene are present in normal cells.

Solution hybridization has been used to measure *HPRT* mRNA levels for a select group of human mutations, including three nonsense and five missense mutations [126]. These nonsense mutations resulted in only a modest decrease in mRNA abundance, 20–46% of normal. Missense mutations resulted in mRNA abundance from 3 to 165% of normal. However, two of the nonsense mutations were in exon 8, which is past the region (Fig. 1) where an effect on mRNA abundance is expected. The one internal nonsense mutation (nucleotide 151) was associated with the largest mRNA abundance (46%). Northern blotting has not been used to assess mRNA abundance for nonsense mutants of the human *HPRT* gene and neither has quantitative RT-PCR been used as have been done for the hamster gene [37,36]. Either the 5′ boundary for a nonsense-mediated effect on mRNA abundance is more internal for the human *HPRT* gene than for the hamster *Hprt* gene (Section 2.3) or nonsense mutations may not have an effect on mRNA abundance for the human *HPRT* gene.

Generally, however, too little amplified cDNA for sequencing is obtained after 1 round of RT-PCR from nonsense mutants for both the *HPRT* gene and the thymidine kinase gene of human lymphoblastoid cells, which is normally adequate for sequencing missense mutants (personal communication, A.J. Grosovsky). Mutations that result in so little mRNA that no cDNA could be amplified after two rounds of PCR include a large proportion (36%) that are nonsense mutations when sequenced from genomic DNA [23]. If a reduction in mRNA abundance for nonsense mutations is typical for internal nonsense mutations in the human gene, then the necessary conditions are present for the selective appearance in nonsense mutants of exon-deleted mRNA species by RT-PCR artifact. Many studies characterizing *hprt* mutants use RT-PCR conditions similar to the preparative protocol for the hamster gene (Section 2.1) [25,127–130]. It is, therefore, a reasonable hypothesis that the appearance of minor amounts of exon 2,3-deleted cDNA from human cells with nonsense mutations is the result of more effective amplification of a normal, minor species.

A sensitive, quantitative analysis of expressed mRNA from human cells will be necessary to establish which of the exon-deleted species are normally produced mRNAs and whether they are increased in abundance by nonsense mutation. This information is needed in order to determine whether the exon-deleted species of *HPRT* mRNA observed in human cells with nonsense mutations arise from the same mechanism as shown for hamster cells.

5.2. Cystic fibrosis gene

The available data for this gene does exhibit a greater reduction of mRNA abundance for nonsense mutations at internal locations than at the ends of the coding sequence, although the reduction in mRNA abundance for mutations at the 3′ end of the gene is variable (Table 3). Furthermore, the quantitative measurements were done with different allele-specific probes and may not be strictly comparable (footnote a, Table 3). The out-of-frame deletion of exon 3 as well as the deletion of exons 3 and 4 produce the first nonsense codon probably far enough into the gene (210 and 171 bases, respectively) to reduce mRNA abundance (*CFTR* entry, Table 3) [84]. However, the out-of-frame deletion of exons 2 and 3 produces its first nonsense codon after 99 bases and may not have a severe effect on mRNA abundance. Therefore, the exon 2,3-deleted species might appear by an RT-PCR artifact, but the other two are less likely by this mechanism because they may also

suffer a reduction of mRNA abundance (Section 2.3).

The mutations causing exon 3 deletions are in either a purine-rich or an A/C-rich region. The exon 20 mutation (598 bases from the 3′ end of the coding sequence) may be past the 3′ boundary for nonsense-mediated effects on mRNA abundance (footnote a, Table 3) since it is associated with normal mRNA abundance. Therefore, it might appear by RT-PCR artifact. However, this mutation is associated with the use of cryptic splice sites in the adjacent exons and is in a purine-rich sequence. Most of the *CFTR* mutations in Table 1 introduce a T near the end of an exon. The one internal mutation, R75X, is in a short purine-rich sequence or an extended A/C-rich sequence. Consequently, these mutations may be affecting splicing enhancers that contribute to exon definition.

5.3. Human dystrophin gene

Patients with the out-of-frame genomic deletion of exons 3–7 of the dystrophin gene have a mild form of dystrophy known as Becker muscular dystrophy, which is normally associated with in-frame deletions [10]. The more severe Duchenne muscular dystrophy usually results from out-of-frame genomic deletions. The milder clinical phenotype of patients with a genomic deletion of exons 3–7 correlates with the presence of minor, in-frame, exon-deleted mRNA (exons 2–7 or 3–9 deleted) [10]. However, the milder phenotype is more likely the result of reinitiation within exon 8 (the first exon after the deletion), producing a normal protein lacking the N-terminal sequence (79 total exons) [92].

The first nonsense codon in dystrophin mRNA with a deletion of exons 3–7 is in exon 8, 130 bp from the initiating AUG and this mRNA has relatively normal abundance (> 25%, Table 3) [85]. Out-of-frame deletions more distal in the gene do cause a reduction in mRNA levels to 10% of normal. This boundary is consistent with other 5′ boundaries for the effect of nonsense mutations on mRNA abundance (Table 3). Since the out-of-frame, major species has normal abundance, the appearance of the minor products is probably not the result of RT-PCR artifact (Sections 2.2 and 2.3). The conflicting results from different workers regarding the presence of the minor, exon-deleted products [92] could reflect the fact that different genomic deletions resulting in the loss of exons 3–7 in dystrophin mRNA probably have different genomic breakpoints, and consequently different pre-mRNA structures. The position of these breakpoints in the mRNA could have a structural effect on exon definition and directly affect the appearance of exon-deleted products.

Minor amounts of in-frame, exon-deleted mRNA have also been observed for the human dystrophin gene in patients with ‘revertant’ fibers: individual muscle fibers that show normal staining to dystrophin antibodies [92,131,132]. Because these fibers appear to be of clonal origin, it is thought that they have acquired a further somatic mutation affecting splicing that removes another exon and restores the reading frame [131,92]. The same revertant fibers have been observed for a mouse model of muscular dystrophy in which in-frame exon-deleted species are present in the affected tissue and also found in small quantity in normal cells [133,134]. Until an analysis is done that separates the revertant fibers from affected fibers before PCR, it will not be possible to positively correlate revertant fibers with exon-skipping.

5.4. Human transacylase gene (E2) of the branched chain α-keto acid dehydrogenase complex

The *E2* gene is formally included by our selection criteria that an exon containing a nonsense mutation was deleted in an mRNA species, but the reported nonsense mutations have the unusual feature that not all exon deletions remove the nonsense mutations. Two nonsense mutations, one in exon 2 and one in exon 6, are both associated with the deletion of exon 4 alone or exon 6 alone. The deletion of exons 2–8 could not be assigned to either mutated allele since they both were in the same patient and the deletion removes both mutations. In addition, the level of the mRNA is not reduced and the proportion of exon-deleted transcripts is relatively high. The data was collected by sequencing 10 cloned cDNAs from amplified cDNA.

The exon 2 mutation is a 2-base deletion in a purine-rich sequence less than 30 bases from the end of exon 2. The exon 6 mutation is a base pair substitution in the middle of a 117 base exon and introduces a T into a purine-rich sequence. If these mutations are causing exon skipping through a struc-

tural effect on enhancers, then mutations in exonic sites can result in the loss of a distant exon without necessitating loss of the exon containing the mutation. Coordinate splicing of several exons is suggested by the exon-deleted species of the *Hprt* gene, where single base pair substitutions can increase the abundance of multiply exon-deleted species [24] (Table 1). Usually, multiple exon deletions affect the exon with the mutation [36,118,24], but some reports involve an adjacent or distant exon (Sections 3.4 and 4.1).

It could be questioned whether the incidence of the exon-deleted species is actually increased by the nonsense mutations, since the increase is based on only ten clones sequenced at random, rather than a bulk biochemical analysis. Some bias during cloning or PCR may have increased the proportion of normal, minor exon-deleted species. Other exon-deleted species were found from normal cells, but not the ones associated with these mutations. These results are perhaps the most atypical of any gene.

6. Conclusions

The three general explanations for the association of exon skipping with nonsense codons are summarized in Table 5. When considering the significance of data correlating nonsense codons with the deletions of exons in mRNA, two of these explanations should first be considered. If RT-PCR has been used, an artifact is possible if nonsense codons reduce the abundance of mRNA for the gene, and if the exon deletion is in-frame. Out-of-frame deletions might also appear through an RT-PCR artifact if the first nonsense codon produced by the exon deletion is near the ends of the coding sequence and the full-length species has the nonsense codon more internal.

Table 5
Explanations for the association between nonsense mutations and exon skipping

Description	Experimental characteristics expected	Example
(1) PCR artifact based on the reduction of abundance of mRNA containing a nonsense codon. When full-length mRNA is abundant, RT-PCR conditions can suppress the amplification of minor mRNA species so that they are not visible on agarose gels. When a nonsense codon reduces the abundance of full-length mRNA, an in-frame exon deletion will maintain normal abundance of the minor species and be more readily amplified. An out-of-frame deletion that moves the first nonsense codon close to the ends of the coding sequence also maintains normal abundance.	(1) Nonsense codons reduce abundance of full-length mRNA. (2) Exon-deleted mRNA species are present in low abundance in normal cells. (3) Exon deletion in minor mRNA species are in-frame **or** the nonsense codon after the frameshift is within about 100 bases of the 5′ end or about 100–400 bases of the 3′ end of the coding sequence. (4) PCR is used to detect the mRNA species.	*Hprt* gene of CHO cells [36] (Section 2)
(2) Structural effect of nonsense mutation on exon definition. Rather than a functional effect of coding potential, the mutated base itself affects exon definition and selection. Presumably, the mutation alters the function of a splicing enhancer in the exonic sequence or a *cis* determinant of exon definition.	(1) Mutation is often within 30 bases of the nearest exon boundary and often introduces a T into a purine-rich sequence. (2) A missense mutation at the same or nearby nucleotide may also cause the same exon deletions in mRNA.	*Mup* gene [44] and *Hprt* gene [36] (Section 3.2, Table 2)
(3) Nuclear scanning—the coding potential of RNA is recognized in the nucleus. Some nuclear entity, yet unknown, recognizes nonsense codons in pre-mRNA or mRNA and results in the skipping or aberrant splicing of those exons containing nonsense or results in enrichment of exon-deleted mRNA.	Nonsense but not missense mutations are associated with exon deletions in mRNA.	Proposed for the *FBN1* gene [55], MVM [102], immunoglobulin κ chain [135], and *TCR-*β gene [89,59]

However, if an abundant major species is exon-deleted, RT-PCR artifact is not a likely explanation. It may not be known for a given gene whether nonsense codons reduce mRNA abundance, but if an RT-PCR artifact is suspected, diluting the input RNA or amplifying PCR products after digestion with a restriction enzyme that cleaves the full-length PCR product may reveal whether exon-deleted products are present in comparable quantity in normal cells. Reducing the number of cycles of PCR and adding radioactive isotope to detect the product may also suffice. However, quantitative Northern blots or ribonuclease protection assays eliminate any concern about PCR artifacts.

The second explanation to consider is that the mutation suppresses proper exon definition, possibly in an exonic splicing enhancer. Base substitutions known to have this effect often introduce a T into a purine-rich sequence. Both missense and nonsense mutations associated with exon skipping are often within 30 bases of the end of an exon, but some internal missense mutations are known. Therefore, any exonic position could be considered as having a structural effect on exon definition. Since exonic splicing enhancers do not have recognizable consensus sequences and because mutations generally are more likely to occur in purine-rich sequences, it is difficult to conclusively attribute exon skipping to an effect on a splicing enhancer. Although this possibility is often discarded after brief consideration because of the lack of simple methods to identify enhancers, more examples of the association of nonsense codons with exon skipping caused by the disruption of exonic enhancers can be expected as more detailed studies are done.

Nuclear scanning is a challenging hypothesis that is still being investigated as a possible mechanism for the association of nonsense mutation and exon skipping, but has not yet been established as the mechanism for any gene. One limitation of the current data is the relatively few nonsense mutations that have been characterized for each gene considered to display nuclear scanning and the even fewer number of missense mutations characterized. Investigation into this possible mechanism may have more impact on the elucidation of the mechanism by which nonsense codons affect mRNA abundance than on exon skipping.

Notes added in proof:

A nonsense mutation in the donor site of exon 5 of the Fanconi Anemia C gene is associated with the skipping of exon 6 (Foe et al., 1998, *Human Mutation* suppl. 1, S25-S27). With this in-frame exon deletion, the mutation is part of a missense codon, which would be expected to maintain the stability of exon-deleted mRNA. No exon 6-deleted mRNA is found in normal cells. This report lends weight to the argument that exonic mutations affecting splice-site choice may act across introns. Also, nonsense mutations in the *NFI* gene that result in exon skipping have been related to changes in proposed mRNA seconday structure, which is not changed for some nearby nonsense codons that do not result in exon skipping (Hoffmeyer et al., 1998, *Am J. Human Gen.*, 62:269-277).

An internal nonsense mutation (exon 24) in the fibrillin gene is associated with normal mRNA abundance (Karttunen et al., 1998, *Human Mutation* suppl. 1, S34-S37). Normal abundance for a nonsense mutation away from the 3′ end of the coding sequence is an exception both for the fibrillin gene and generally. A rule for the 3′ boundary of a nonsense-mediated effect on mRNA abundance states that mRNA abundance will be reduced if the nonsense codon is more than 50-55 nucleotides upstream from the last exon/exon boundary (Nagy and Maquat, 1998, *Trends in Biochemical Sciences*, 23: 198–199). The TCR-β gene is also an exception to this rule in that normal mRNA abundance is found up to 8 nucleotides from the exon/exon boundary [89,59].

Acknowledgements

I am grateful to Miles F. Wilkinson for critical revision of the manuscript before submission. I am also thankful for the expert comments of the reviewers of The Journal, which contributed to improvements in the manuscript.

References

[1] H.C. Dietz, D. Valle, C.A. Francomano, R.J. Kendzior Jr., R.E. Pyerite, G.R. Cutting, The skipping of constitutive exons in vivo induced by nonsense mutations, Science 259 (1993) 680–683.

[2] I. Santisteban, F.X. Arredondo-Vega, S. Kelly, M. Loubser, N. Meydan, C. Roifman, P.L. Howell, R. Bowen, K. Weinberg, M.L. Schroeder, M.S. Hershfield, Three new adenosine deaminase mutations that define a splicing enhancer and cause severe and partial phenotypes: implications for evolution of a CpG hotspot and expression of a transduced ADA cDNA, Hum. Mol. Genet. 4 (1995) 2081–2087.

[3] P. Jarolim, H.L. Rubin, V. Brabec, J. Palek, A nonsense mutation 1669Glu → Ter within the regulatory domain of human erythroid ankyrin leads to a selective deficiency of the major ankyrin isoform (band 2.1) and a phenotype of autosomal dominant hereditary spherocytosis, J. Clin. Invest. 95 (1995) 941–947.

[4] J. Hull, S. Shackleton, A. Harris, The stop mutation R553X in the *CFTR* gene results in exon skipping, Genomics 19 (1994) 362–717.

[5] K. Will, T. Dörk, M. Stuhrmann, H. von der Hardt, H. Ellemunter, B. Tümmler, J. Schmidtke, Transcript analysis of *CFTR* nonsense mutations in lymphocytes and nasal epithelial cells from cystic fibrosis patients, Hum. Mutat. 5 (1995) 210–220.

[6] T. Bienvenu, C. Beldjord, J. Chelly, N. Fonknechten, D. Hubert, D. Dusser, J.C. Kaplan, Analysis of alternative splicing patterns in the cystic fibrosis transmembrane conductance gene using mRNA derived from lymphoblastoid cells of cystic fibrosis patients, Eur. J. Hum. Genet. 4 (1996) 127–134.

[7] J. Zielenski, R. Rozmahel, D. Bozon, B. Kerem, Z. Grzelczak, J.R. Riordan, J. Rommens, L.C. Tsui, Genomic sequence of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, Genomics 10 (1991) 229–235.

[8] B.-S. Kerem, J. Zielenski, D. Markiewicz, D. Bozon, E. Gazit, J. Yahav, D. Kennedy, J.R. Riordan, F.S. Collins, J.M. Rommens, L.-C. Tsui, Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 8447–8451.

[9] N. Shiga, Y. Takeshima, H. Sakamoto, K. Inoue, Y. Yokota, M. Yokoyama, M. Matsuo, Disruption of the splicing enhancer sequence within exon 27 of the dystrophin gene by a nonsense mutation induces partial skipping of the exon and is responsible for Becker muscular dystrophy, J. Clin. Invest. 100 (1997) 2204–2210.

[10] J. Chelly, H. Gilgenkrantz, M. Lambert, G. Hamard, P. Chafey, D. Récan, P. Katz, A. de la Chapelle, M. Koenig, L.B. Ginjaar, M. Fardeau, F. Tomé, A. Kahn, J.-C. Kaplan, Effect of dystrophin gene deletions on mRNA levels and processing in Duchenne and Becker muscular dystrophies, Cell 63 (1990) 1239–1248.

[11] C.W. Fisher, C.R. Fisher, J.L. Chuang, K.S. Lau, D.T. Chuang, R.P. Cox, Occurrence of a 2-bp (AT) deletion allele and a nonsense (G-to-T) mutant allele at the E2 (DBT) locus of six patients with maple syrup urine disease: multiple-exon skipping as a secondary effect of the mutations, Am. J. Hum. Genet. 52 (1993) 414–424.

[12] K.S. Lau, J.L. Chuang, W.J. Herring, D.J. Danner, R.P. Cox, D.R. Chuang, The complete cDNA sequence for dihydrolipoyl transacylase (E2) of human branched-chain α-keto acid dehydrogenase complex, Biochim. Biophys. Acta 1132 (1992) 319–321.

[13] K.S. Lau, W.J. Herring, J.L. Chuang, M. McKean, D.J. Danner, R.P. Cox, D.T. Chuang, Structure of the gene encoding dihydrolipoyl transacylase (*E2*) component of human branched chain α-keto acid dehydrogenase complex and characterization of an *E2* pseudogene, J. Biol. Chem. 267 (1992) 24090–24096.

[14] J.A. Naylor, P.M. Green, C.R. Rizza, F. Giannelli, Analysis of factor VIII mRNA reveals defects in everyone of 28 haemophilia A patients, Hum. Mol. Genet. 2 (1993) 11–17.

[15] W.I. Wood, D.J. Capon, C.C. Simonsen, D.L. Eaton, J. Gitschier, B. Keyt, P.H. Seeburg, D.H. Smith, P. Hollingshead, K.L. Wion, E. Delwart, E.G.D. Tuddenham, G.A. Vehar, R.M. Lawn, Expression of active human factor VIII from recombinant DNA clones, Nature 312 (1984) 326–330.

[16] R.A. Gibson, A. Hajianpour, M. Murer-Orlando, M. Buchwald, C.G. Mathew, A nonsense mutation and exon skipping in the Fanconi anaemia group C gene, Hum. Mol. Genet. 2 (1993) 797–799.

[17] C.A. Strathdee, H. Gavish, W.R. Shannon, M. Buchwald, Cloning of cDNAs for Fanconi's anaemia by functional complementation, correction, Nature 358 (1992) 434.

[18] J. Pié, N. Casals, C.H. Casale, C. Buesa, C. Mascaro, A. Barcelo, M.-O. Rolland, T. Zabot, D. Haro, F. Eyskens, P. Divry, F.G. Hegardt, A nonsense mutation in the 3-hydroxy-3-methylglutaryl-CoA lyase gene produces exon skipping in two patients of different origin with 3-hydroxy-3-methylglutaryl-CoA lyase deficiency, Biochem. J. 323 (1997) 329–335.

[19] C.H. Casale, N. Casals, J. Pié, N. Zapater, C. Pérez-Cerdá, B. Meriner, M. Martínez-Pardo, J.J. García-Peñas, J.M. García-Gonzalez, R. Lama, B.-T. Poll-The, J.A.M. Smeitink, R.J.A. Wanders, M. Ugarte, F.G. Hegardt, A nonsense mutation in the exon 2 of the 3-hydroxy-3-methylglutaryl coenzyme A lyase (*HL*) gene producing three mature mRNAs is the main cause of 3-hydroxy-3-methylglutaric aciduria in European Mediterranean patients, Arch. Biochem. Biophys. 349 (1998) 129–137.

[20] F. Endo, H. Awata, H. Katoh, I. Matsuda, A nonsense mutation in the 4-hydroxyphenylpyruvic acid dioxygenase gene (*Hpd*) causes skipping of the constitutive exon and hypertyrosinemia in mouse strain III, Genomics 25 (1995) 164–169.

[21] G.G. Jolly, J.E. Shively, T. Hunkapillar, T. Friedmann, Isolation and characterization of a full-length expressible cDNA for human hypoxanthine phosphoribosyl transferase, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 477–481.

[22] O. Rigaud, A. Laquerbe, E. Moustacchi, DNA sequence analysis of *hprt* mutants induced in human lymphoblastoid cells adapted to ionizing radiation, Radiat. Res. 144 (1995) 181–189.

[23] M. Khaidakov, D. Young, H. Erfle, A. Mortimer, Y. Voronkov, B.W. Glickman, Molecular analysis of mutations in T-lymphocytes from experienced Soviet cosmonauts, Environ. Mol. Mutagen. 30 (1997) 21–30.

[24] A.-M. Österholm, S.-M. Hou, Splice mutations at the *hprt* locus in human T-lymphocytes, Environ. Mol. Mutagen. 24 (1998) in press.
[25] H. Steingrimsdottir, G. Rowley, G. Dorado, J. Cole, A.R. Lehmann, Mutations which alter splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene, Nucleic Acids Res. 20 (1992) 1201–1208.
[26] J.-L. Yang, M.-C. Hu, C.-W. Wu, Novel mutational spectrum induced by *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine in the coding region of the hypoxanthine (guanine) phosphoribosyltransferase gene in diploid human fibroblasts, J. Mol. Biol. 221 (1991) 421–430.
[27] S.L. Nelson, C.R. Giver, A.J. Grosovsky, Spectrum of X-ray-induced mutations in the human *hprt* gene, Carcinogenesis 15 (1994) 495–502.
[28] P.-M.L. Morgenthaler, D. Holzhäuser, Analysis of mutations induced by 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) in human lymphoblastoid cells, Carcinogenesis 16 (1995) 713–718.
[29] J.-L. Yang, P.-C. Lee, S.-R. Lin, J.-G. Lin, Comparison of mutation spectra induced by *N*-ethyl-*N*-nitrosourea in the *hprt* gene of Mer^+ and Mer^- diploid human fibroblasts, Carcinogenesis 15 (1994) 939–945.
[30] D. Papadopoulo, A. Laquerbe, C. Guillouf, E. Moustacchi, Molecular spectrum of mutations induced at the *HPRT* locus by a cross-linking agent in human cell lines with different repair capacities, Mutat. Res. 294 (1993) 167–177.
[31] J.-L. Yang, J.G. Lin, M.-C. Hu, C.-W. Wu, The *hprt* gene in G_1 and late S phase of diploid human fibroblasts, Cancer Res. 53 (1993) 2865–2873.
[32] D.S. Konecki, J. Brennand, J.C. Fuscoe, C.T. Caskey, A.C. Chinault, Hypoxanthine-guanine phosphoribosyltransferase genes of mouse and Chinese hamster: construction and sequence analysis of cDNA recombinants, Nucleic Acids Res. 10 (1982) 6763–6775.
[33] E.E. Hennig, A.H. Conney, S.-J.C. Wei, Characterization of *hprt* splicing mutations induced by the ultimate carcinogenic metabolite of benzo[α]pyrene in Chinese hamster V-79 cells, Cancer Res. 55 (1995) 1550–1558.
[34] C.R. Valentine, R.H. Heflich, Genomic DNA sequencing of mRNA splicing mutants in the *hprt* gene of Chinese hamster ovary cells, Environ. Mol. Mutagen. 25 (1995) 85–96.
[35] P.A. Estes, N.E. Cookes, S.A. Liebhaber, A native RNA secondary structure controls alternative splice-site selection and generates two human growth hormone isoforms, J. Biol. Chem. 267 (1992) 14902–14908.
[36] C.R. Valentine, R.H. Heflich, The association of nonsense mutation with exon skipping in *Hprt* mRNA of Chinese hamster ovary cells results from an artifact of RT-PCR, RNA 3 (1997) 660–676.
[37] M.G. Manjanatha, L.A. Lindsey, R.A. Mittelstaedt, R.H. Heflich, Low *hprt* levels and multiple *hprt* mRNA species in 6-thioguanine-resistant Chinese hamster cell mutants possessing nonsense mutations, Mutat. Res. 308 (1994) 65–75.
[38] P. Menichini, A. Inga, G. Fronza, R. Iannone, P. Degan, P. Campomenosi, A. Abbondandolo, Defective splicing induced by 4NQO in the hamster *hprt* gene, Mutat. Res. 323 (1994) 159–165.
[39] L.-H. Zhang, D. Jenssen, Site specificity of *N*-methyl-*N*-nitrosourea-induced transition mutations in the *hprt* gene, Carcinogenesis 12 (1991) 1903–1909.
[40] G. Bach, S.M. Moskowitz, P.T. Tieu, A. Matynia, E.F. Neufeld, Molecular analysis of Hurler syndrome in Druze and Muslim Arab patients in Israel: multiple allelic mutations of the *IDUA* gene in a small geographic area, Am. J. Hum. Genet. 53 (1993) 330–338.
[41] H.S. Scott, D.S. Anson, A.M. Orsborn, P.V. Nelson, P.R. Clements, C.P. Morris, J.J. Hopwood, Human α-L-iduronidase: cDNA isolation and expression, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 9695–9699.
[42] S. Markiewicz, A. Subtil, A. Dautry-Varsat, A. Fischer, G. de Saint Basile, Detection of three nonsense mutations and one missense mutation in the interleukin-2 receptor γ chain gene in SCIDX1 that differently affect the mRNA processing, Genomics 21 (1994) 291–293.
[43] T. Takeshita, H. Asao, K. Ohtani, N. Ishii, S. Kumaki, N. Tanaka, H. Munakata, M. Nakamura, K. Sugamura, Cloning of the γ chain of the human IL-2 receptor, Science 257 (1992) 379–382.
[44] P. Belgrader, L.E. Maquat, Nonsense but not missense mutations can decrease the abundance of nuclear mRNA for the mouse major urinary protein, while both types of mutations can facilitate exon skipping, Mol. Cell. Biol. 14 (1994) 6326–6336.
[45] S. Das, B. Levinson, S. Whitney, C. Vulpe, S. Packman, J. Gitshier, Diverse mutations in patients with Menkes disease often lead to exon skipping, Am. J. Hum. Genet. 55 (1994) 883–889.
[46] L. Messiaen, T. Callens, A. DePaepe, M. Craen, G. Mortier, Characterization of two different nonsense mutations, C6972A and C6792G, causing skipping of exon 37 in the *NF1* gene, Hum. Genet. 101 (1997) 75–80.
[47] R. Okamoto, T. Yamazaki, A. Katsumi, T. Kojima, J. Takamatsu, M. Nishida, H. Saito, A novel nonsense mutation associated with an exon skipping in a patient with hereditary protein S deficiency type I, Thromb. Haemost. 75 (1996) 877–882.
[48] D.K. Schmidel, A.V. Tatro, L.G. Phelps, J.A. Tomczak, G.L. Long, Organization of the human protein S genes, Biochemistry 29 (1990) 7845–7852.
[49] M.H. Ricketts, M.J. Simons, J. Parma, L. Mercken, Q. Dong, G. Vassart, A nonsense mutation causes hereditary goitre in the Afrikander cattle and unmasks alternative splicing of thyroglobulin transcripts, Proc. Natl. Acad. Sci. U.S.A. 84 (1987) 3181–3184.
[50] H.M. Targovnik, G. Medeiros-Neto, V. Varela, P. Cochaux, B.L. Wajchenberg, G. Vassart, A nonsense mutation causes human hereditary congenital goiter with preferential production of a 171-nucleotide-deleted thyroglobulin ribonucleic acid messenger, J. Clin. Endocrinol. Metab. 77 (1993) 210–215.
[51] A. Zhu, M. Zhang, R.M. Blaese, J.M.J. Derry, A. Junker, U. Francke, S.-H. Chen, H.D. Ochs, The Wiskott–Aldrich syndrome and X-linked congenital thrombocytopenia are caused by mutations of the same gene, Blood 86 (1995) 3797–3804.

[52] J.M. Derry, H.D. Ochs, U. Francke, Isolation of a novel gene mutated in Wiskott–Aldrich syndrome, Cell 78 (1994) 635–644.
[53] A. Hamosh, B.J. Rosenstein, G.R. Cutting, CFTR nonsense mutations G542X and W1282X associated with severe reduction of CFTR mRNA in nasal epithelial cells, Hum. Mol. Genet. 1 (1992) 542–544.
[54] T. Shoshani, E. Kerem, A. Szeinberg, A. Augarten, Y. Yahav, D. Cohen, J. Rivlin, A. Tal, B.s. Kerem, Similar levels of mRNA from the W1282X and the ΔF508 cystic fibrosis alleles, in nasal epithelial cells, J. Clin. Invest. 93 (1994) 1502–1507.
[55] H.C. Dietz, R.J. Kendzior Jr., Maintenance of an open reading frame as an additional level of scrutiny during splice site selection, Nat. Genet. 8 (1994) 183–188.
[56] K. Tanaka, A. Watakabe, Y. Shimura, Polypurine sequences within a downstream exon function as a splicing enhancer, Mol. Cell. Biol. 14 (1994) 1347–1354.
[57] L.E. Maquat, Defects in RNA splicing and the consequence of shortened translational reading frames, Am. J. Hum. Genet. 59 (1996) 279–286.
[58] L.E. Maquat, When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells, RNA 1 (1995) 453–465.
[59] S. Li, M.F. Wilkinson, Nonsense surveillance in lymphocytes, Immunity 8 (1998) 135–141.
[60] R.K. Newton, R.A. Mittelstaedt, R.H. Heflich, Analysis of solvent control and 1-nitrosopyrene-induced Chinese hamster ovary cell mutants by Southern and Northern blots and the polymerase chain reaction, Environ. Mol. Mutagen. 19 (1992) 147–155.
[61] R.K. Newton, R.A. Mittelstaedt, M.G. Manjanatha, R.H. Heflich, DNA sequence analysis of 1-nitrosopyrene-induced mutations in the *hprt* gene of Chinese hamster ovary cells, Carcinogenesis 13 (1992) 819–825.
[62] H. Ozsahin, F.X. Arredondo-Vegas, I. Santisteban, H. Fuhrer, P. Tuchschmidt, W. Jochum, A. Aguzzi, H.M. Lederman, A. Fleischman, J.A. Winkelstein, R.A. Seger, M.S. Hershfield, Adenosine deaminase deficiency in adults, Blood 89 (1997) 2849–2855.
[63] D.A. Wiginton, D.J. Kaplan, J.C. States, A.L. Akeson, C.M. Perme, I.J. Bilyk, A.J. Vaughn, D.L. Lattier, J.J. Hutton, Complete sequence and structure of the gene for human adenosine deaminase, Biochemistry 25 (1986) 8234–8244.
[64] W. Liu, C. Qian, U. Francke, Silent mutation induces exon skipping of fibrillin-1 gene in Marfan syndrome, Nat. Genet. 16 (1997) 238–329.
[65] Y. Jin, H.C. Dietz, R.A. Montgomery, W.R. Bell, I. McIntosh, B. Coller, P.F. Bray, Glanzmann thrombasthenia—cooperation between sequence variants in *cis* during splice site selection, J. Clin. Invest. 98 (1996) 1745–1754.
[66] N. Wakamatsu, H. Kobayashi, T. Myatake, S. Tsuji, A novel exon mutation in the human β-hexosaminidase β subunit gene affects 3′ splice site selection, J. Biol. Chem. 267 (1992) 2406–2413.
[67] K.J. Burkhart-Schultz, C.L. Thompson, I.M. Jones, Spectrum of somatic mutation at the hypoxanthine phosphoribosyltransferase (*hprt*) gene of healthy people, Carcinogenesis 17 (1996) 1871–1883.
[68] C. Curry, G.T. Rowley, V. Saddi, D. Beare, J. Cole, B.W. Glickman, Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived in vivo T-lymphocyte mutants, Environ. Mol. Mutagen. 25 (1995) 169–179.
[69] L. Recio, D. Simpson, J. Cochrane, H. Liber, T.R. Skopek, Molecular analysis of *hprt* mutants induced by 2-cyanoethylene oxide in human lymphoblastoid cells, Mutat. Res. 242 (1990) 195–208.
[70] E. Darè, L.-H. Zhang, D. Jenssen, V. Bianchi, Molecular analysis of mutations in the *Hprt* gene of V79 hamster fibroblasts: effects of imbalances in the dCTP, dGTP, and dTTP pools, J. Mol. Biol. 252 (1995) 514–521.
[71] L.-H. Zhang, H. Vrieling, A.A. van Zeeland, D. Jenssen, Spectrum of spontaneously occurring mutations in the *hprt* gene of V79 Chinese hamster cells, J. Mol. Biol. 223 (1992) 627–635.
[72] D.H. Llewellyn, G.A. Scobie, A.J. Urquhart, S.D. Whatley, A.G. Roberts, P.R. Harrison, G.H. Elder, Acute intermittent porphyria caused by defective splicing of porphobilinogen deaminase RNA: a synonymous codon mutation at −22 bp from the 5′ splice site causes skipping of exon 3, J. Med. Genet. (1996) 437–438.
[73] L. De Meirleir, W. Lissens, C. Benelli, G. Ponsot, I. Gesguerrre, C. Marasc, D. Rodriguez, J.-M. Saudbray, F. Poggi, I. Liebaers, Aberrant splicing of exon 6 in the pyruvate dehydrogenase-E1α mRNA linked to a silent mutation in a large family with Leigh's encephalomyelopathy, Pediatr. Res. 36 (1994) 707–712.
[74] J. Ross, mRNA stability in mammalian cells, Microbiol. Rev. 59 (1995) 423–450.
[75] A. Jacobson, S.W. Peltz, Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells, Ann. Rev. Biochem. 65 (1996) 693–739.
[76] G. Caponigro, R. Parker R, Mechanisms and control of mRNA turnover in *Saccharomyces cerevisiae*, Microbiol. Rev. 60 (1996) 233–249.
[77] L.A. Chasin, G. Urlaub, P. Mitchell, C. Ciudad, J. Barth, A.M. Carothers, R. Steigerwalt, D. Grunberger, RNA processing mutants at the dihydrofolate reductase locus in Chinese hamster ovary cells, in: M.L. Mendelsohn, R.J. Albertini (Eds.), Mutation and the Environment, Part A, Basic Mechanisms, Wiley-Liss, New York, 1990, pp. 295–304.
[78] P. Belgrader, J. Cheng, X. Zhou, L.S. Stephenson, L.E. Maquat, Mammalian nonsense codons can be *cis* effectors of nuclear mRNA half-life, Mol. Cell. Biol. 14 (1994) 8219–8228.
[79] B. Baumann, M.J. Potash, G. Köhler, Consequences of frameshift mutations at the immunoglobulin heavy chain locus of the mouse, EMBO J. 4 (1985) 351–359.
[80] I.-T. Chen, L.A. Chasin, Direct selection for mutations affecting specific splice sites in a hamster dihydrofolate reductase minigene, Mol. Cell. Biol. 13 (1993) 289–300.

[81] F. Mathieu-Daudé, J. Welsh, T. Vogt, M. McClelland, DNA rehybridization during PCR: the 'C_ot effect' and its consequences, Nucleic Acids Res. 11 (1996) 2080–2086.
[82] I. Vořechovský, L. Luo, J.M. Hertz, S.S. Frøland, R. Klemola, M. Fiorini, I. Quinti, R. Paganelli, H. Ozsahin, L. Hammarström, A.D.B. Webster, C.I.E. Smith, Mutation pattern in the Bruton's tyrosine kinase gene in 26 unrelated patients with X-linked agammaglobulinemia, Hum. Mutat. 9 (1997) 418–425.
[83] D. Vetrie, I. Vořechovský, P. Sideras, J. Holland, A. Davies, F. Flinter, L. Hammarström, C. Kinnon, R. Levinsky, M. Bobrow, C.I.E. Smith, D.R. Bentley, The gene involved in X-linked agammaglobulinaemia is a member of the *src* family of protein-tyrosine kinases, Nature 361 (1993) 226–233.
[84] J.R. Riordan, J.M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J.-L. Chou, M.L. Drumm, M.C. Iannuzzi, F.S. Collins, L.-C. Tsui, Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA, Science 245 (1989) 1066–1073.
[85] M. Koenig, A.P. Monaco, L.M. Kunkel, The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein, Science 238 (1988) 347–350.
[86] H.C. Dietz, R.E. Pyeritz, Mutations in the human gene for fibrillin-1 (*FBN*1) in the Marfan syndrome and related disorders, Hum. Mol. Genet. 4 (1995) 1799–1809.
[87] L. Pereira, M. D'Alessio, F. Ramirez, J.R. Lynch, B. Sykes, T. Pangilinan, J. Bonadio, Genomic organization of the sequence coding for fibrillin, the defective gene product in Marfan syndrome, Hum. Mol. Genet. 2 (1993) 961–968.
[88] A. van Hoof, P.J. Green, Premature nonsense codons decrease the stability of phytohemagglutinin mRNA in a position-dependent manner, Plant J. 10 (1996) 415–424.
[89] M.S. Carter, L. Shulin, M.F. Wilkinson, A splicing-dependent regulatory mechanism that detects translation signals, EMBO J. 15 (1996) 5965–5975.
[90] J. Zhang, L.E. Maquat, Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells, EMBO J. 16 (1997) 826–833.
[91] J. Cheng, P. Belgrader, X. Zhou, L.E. Maquat, Introns are *cis* effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance, Mol. Cell. Biol. 14 (1994) 6317–6325.
[92] A.V. Winnard, J.R. Mendell, T.W. Prior, J. Florence, A.H.M. Burghes, Frameshift deletions of exons 3–7 and revertant fibers in Duchenne muscular dystrophy: mechanisms of dystrophin production, Am. J. Hum. Genet. 56 (1995) 158–166.
[93] J. Cheng, M. Fogel-Petrovic, L.E. Maquat, Translation to near the distal end of the penultimate exon is required for normal levels of spliced triosephosphate isomerase mRNA, Mol. Cell. Biol. 10 (1990) 5215–5225.
[94] O. Kessler, L.A. Chasin, Effects of nonsense mutations on nuclear and cytoplasmic adenine phosphoribosyltransferase RNA, Mol. Cell. Biol. 16 (1996) 4426–4435.
[95] G. Urlaub, P.J. Mitchell, C.J. Ciudad, L.A. Chasin, Nonsense mutations in the dihydrofolate reductase gene affect RNA processing, Mol. Cell. Biol. 9 (1989) 2868–2880.
[96] S.J. Baserga, E.J. Benz Jr., β-Globin nonsense mutation: deficient accumulation of mRNA occurs despite normal cytoplasmic stability, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 2935–2939.
[97] M.C. Willing, S.P. Deschenes, R. Slayton, E.J. Roberts, Premature chain termination is a unifying mechanism for COL1A1 null alleles in osteogenesis imperfecta type I cell strains, Am. J. Hum. Genet. 59 (1996) 799–809.
[98] P. Belgrader, J. Chen, L.E. Maquat, Evidence to implicate translation by ribosomes in the mechanism by which nonsense codons reduce the nuclear level of human triosephosphate isomerase mRNA, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 482–486.
[99] M.S. Carter, J. Doskow, P. Morris, S. Li, R.P. Nhim, S. Sandstedt, M.F. Wilkinson, A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro, J. Biol. Chem. 270 (1995) 28995–29003.
[100] S.L. Li, D. Leonard, M.F. Wilkinson, T cell receptor (TCR) mini-gene mRNA expression regulated by nonsense codons —a nuclear-associated translation-like mechanism, J. Exp. Med. 185 (1997) 985–992.
[101] S. Aoufouchi, J. Yélamos, C. Milstein, Nonsense mutations inhibit RNA splicing in a cell-free system: recognition of mutant codon is independent of protein synthesis, Cell 85 (1996) 415–422.
[102] L.K. Naeger, R.V. Schoborg, Q. Zhao, G.E. Tullis, D.J. Pintel, Nonsense mutations inhibit splicing of MVM RNA in *cis* when they interrupt the reading frame of either exon of the final spliced product, Genes Dev. 6 (1992) 1107–1119.
[103] C. Andreuttizaugg, R.J. Scott, R. Iggo, Inhibition of nonsense-mediated messenger RNA decay in clinical samples facilitates detection of human MSH2 mutations with an in vivo fusion protein assay and conventional techniques, Cancer Res. 57 (1997) 3288–3293.
[104] T.N. Darling, J.A. Mcgrath, C. Yee, B. Gatalica, R. Hametner, J.W. Bauer, G. Pohlagubo, A.M. Christiano, J. Uitto, H. Hintner, K.B. Yancey, Premature termination codons are present on both alleles of the bullous pemphigoid antigen 2 type XVII collagen gene in five Austrian families with generalized atrophic benign epidermolysis bullosa, J. Invest. Dermatol. 108 (1997) 463–468.
[105] T.N. Darling, C. Yee, B. Koh, J.A. Mcgrath, J.W. Bauer, J. Uitto, H. Hintner, K.B. Yancey, Cyclohexamide facilitates the identification of aberrant transcripts resulting from a novel splice-site mutation in COL17A1 in a patient with generalized atrophic benign epidermolysis bullosa, J. Invest. Dermatol. 110 (1998) 165–169.
[106] S.M. Berget, Exon recognition in vertebrate splicing, J. Biol. Chem. 270 (1995) 2411–2414.
[107] R. Reed, T. Maniatis, A role for exon sequences and splice-site proximity in splice-site selection, Cell 46 (1986) 681–690.
[108] T. Achsel, Y. Shimura, Factors involved in the activation of

pre-mRNA splicing from downstream splicing enhancers, J. Biochem. 120 (1996) 53–60.
[109] L. Coulter, M.A. Landree, T.A. Cooper, Identification of a new class of exonic splicing enhancer by in vivo selection, Mol. Cell. Biol. 17 (1997) 2143–2150.
[110] Z. Wang, H.M. Hoffman, P.J. Grabowski, Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity, RNA 1 (1995) 21–35.
[111] H. Tian, R. Kole, Selection of novel exon recognition elements from a pool of random sequences, Mol. Cell. Biol. 15 (1995) 6291–6298.
[112] T.A. Cooper, W. Mattox, The regulation of splice-site selection, and its role in human disease, Am. J. Hum. Gen. 61 (1997) 259–266.
[113] Q. Sun, A. Mayeda, R.K. Hampson, A.R. Krainer, F.M. Rottman, General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer, Genes Dev. 7 (1993) 2598–2608.
[114] J. Ramchatesingh, A.M. Zahler, K.M. Neugebauer, M.B. Roth, R.A. Cooper, A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer, Mol. Cell. Biol. 15 (1995) 4898–4907.
[115] F. Del Gatto, M.-C. Gesnel, R. Breathnach, The exon sequence TAGG can inhibit splicing, Nucleic Acids Res. 24 (1996) 2017–2021.
[116] J. Cole, T.R. Skopek, Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo, Mutat. Res. 304 (1994) 33–105.
[117] N.F. Cariello, T.R. Skopek, Analysis of mutations occurring at the human *hprt* locus, J. Mol. Biol. 231 (1993) 41–57.
[118] J.P. O'Neill, P.K. Rogan, N. Cariello, J.A. Nicklas, Mutations that alter RNA splicing of the human *HPRT* gene: a review of the spectrum. Rev. Mutat. Res. (1997), in press.
[119] M.J.L. Ligtenberg, A.M.C. Gennissen, H.S. Vos, J. Hilkens, A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA, Nucleic Acids Res. 19 (1991) 297–301.
[120] B. Marshall, G. Isidro, M.G. Boavida, Naturally occurring splicing variants of the *hMSH2* gene containing nonsense codons identify possible mRNA instability motifs within the gene coding region, Biochim. Biophys. Acta 1308 (1996) 88–92.
[121] H.C. Dietz, Nonsense mutations and altered splice-site selection, Am. J. Hum. Genet. 60 (1997) 729–730.
[122] P.M. Smooker, J. Christodoulou, R.R. McInnes, R.G.H. Cotton, A mutation causing *DHPR* deficiency results in a frameshift and a secondary splicing defect, J. Med. Genet. 32 (1995) 220–223.
[123] W. Liu, C. Qian, K. Comeau, T. Brenn, H. Furthmayr, U. Francke, Mutant fibrillin-1 monomers lacking EGF-like domains disrupt microfibril assembly and cause severe Marfan syndrome, Hum. Mol. Genet. 5 (1996) 1581–1587.
[124] I.O. Daar, L.E. Maquat, Premature translation termination mediates triosephosphate isomerase mRNA degradation, Mol. Cell. Biol. 8 (1988) 812–813.
[125] Q. Zhao, R.V. Schoborg, D.J. Pintel, Alternative splicing of pre-mRNAs encoding the nonstructural proteins of minute virus of mice is facilitated by sequences within the downstream intron, J. Virol. 68 (1994) 2849–2859.
[126] A.-M. Steen, S. Sahlén, S.-M. Hou, B. Lambert, *Hprt* activities and RNA phenotypes in 6-thioguanine resistant human T-lymphocytes, Mutat. Res. 286 (1993) 209–215.
[127] B. Andersson, S.-M. Hou, B. Lambert, Mutations causing defective splicing in the human *hprt* gene, Environ. Mol. Mutagen. 20 (1992) 89–95.
[128] B. Andersson, S. Fält, B. Lambert, Strand specificity for mutations induced by (+)-anti BPDE in the *hprt* gene in human T-lymphocytes, Mutat. Res. 269 (1992) 129–140.
[129] S.-M. Hou, A.-M. Steen, S. Fält, B. Andersson, Molecular spectrum of background mutation at the *hprt* locus in human T-lymphocytes, Mutagenesis 8 (1993) 43–49.
[130] J.G. Jansen, G.R. Mohn, H. Vrieling, C.M. van Teijlingen, P.H.M. Lohman, A.A. van Zeeland, Molecular analysis of *Hprt* gene mutations in skin fibroblasts of rats exposed in vivo to *N*-methyl-*N*-nitrosourea or *N*-ethyl-*N*-nitrosourea, Cancer Res. 54 (1994) 2478–2485.
[131] L.T. Thanh, N. thi Man, T.R. Helliwell, G.E. Morris, Characterization of revertant muscle fibers in Duchenne muscular dystrophy, using exon-specific monoclonal antibodies against dystrophin, Am. J. Hum. Genet. 56 (1995) 725–731.
[132] T.G. Sherratt, T. Vulliamy, V. Dubowitz, C.A. Sewry, P.R. Strong, Exon skipping and translation in patients with frameshift deletions in the dystrophin gene, Am. J. Hum. Genet. 53 (1993) 1007–1015.
[133] S.D. Wilton, D.E. Dye, N.G. Laing, Dystrophin gene transcripts skipping the mdx mutation, Muscle Nerve 20 (1997) 728–734.
[134] S.D. Wilton, D.E. Dye, L.M. Blechynden, N.G. Laing, Revertant fibres—a possible genetic therapy for Duchenne muscular dystrophy, Neuromuscular Disord. 7 (1997) 329–335.
[135] S. Aoufouchi, J. Yélamos, C. Milstein, Nonsense mutations inhibit RNA splicing in a cell-free system: recognition of mutant codon is independent of protein synthesis, Cell 85 (1996) 415–422.