

**Pre-mRNA splicing aberrations and cancer**

**Christopher A Pettigrew<sup>1</sup>, Melissa A Brown<sup>1</sup>**

<sup>1</sup>*School of Molecular and Microbial Sciences, University of Queensland, St. Lucia, Queensland, Australia, 4072*

**TABLE OF CONTENTS**

1. Abstract
2. Regulation of splicing
  - 2.1. Splicing elements
    - 2.1.1. Splice site consensus sequences
    - 2.1.2. Exonic splice elements
  - 2.2. Splicing factors
    - 2.2.1. Spliceosome
    - 2.2.2. SR proteins
    - 2.2.3. hnRNPs
3. Defects in splicing of cancer-associated genes
  - 3.1. Mutations in splicing elements
    - 3.1.1. Mutations in consensus splice sites
    - 3.1.2. Mutations that disrupt ESEs
    - 3.1.3. Mutations that create or disrupt other splicing elements
  - 3.2. Dysregulated expression of splicing factors
  - 3.3. Dysregulated expression of splice variants
4. Clinical implications of splicing defects in cancer
  - 4.1. Tumour biomarkers
  - 4.2. Therapeutic strategies for correcting splicing defects
5. Conclusions and perspectives
6. Acknowledgements
7. References

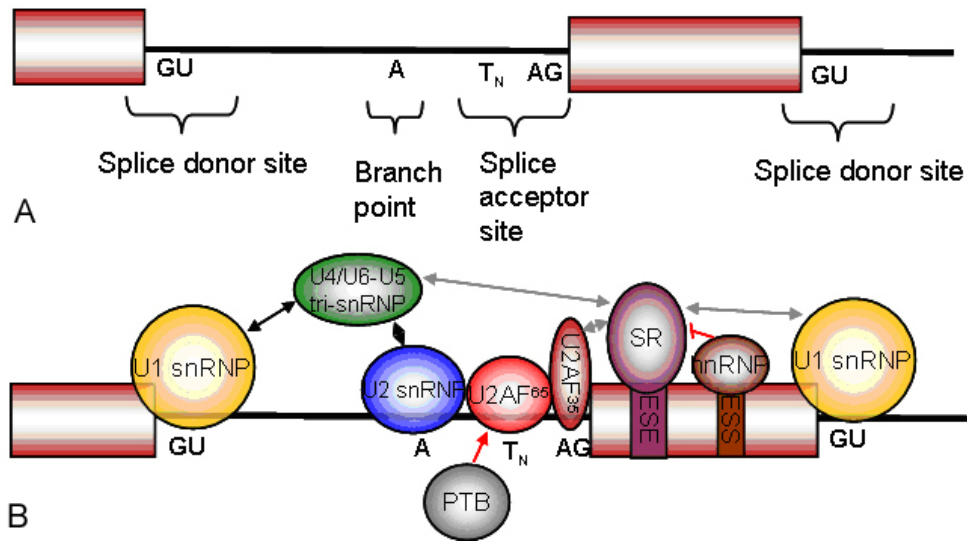
**1. ABSTRACT**

Splicing requires the accurate recognition of exonic sequences from the surrounding thousands of nucleotides of intronic sequence and is achieved by the coordinate interplay of splicing regulatory elements in genes and the *trans*-acting RNA and protein molecules to which they bind. Infidelity in this process can have dramatic consequences for protein production, with errors resulting in mRNA instability or the production of aberrant protein products. It is therefore not surprising that disruptions of splicing processes have been associated with a wide range of diseases, including cancer. This review looks at some of the mechanisms that regulate splicing and how disruption of such mechanisms can contribute to cancer susceptibility and progression.

**2. REGULATION OF SPLICING**

**2.1. Splicing elements**

Accurate excision of intronic sequences requires several sequence elements in the pre-mRNA including a 5' splice site (5' ss), branch point motif, poly-pyrimidine tract and 3' splice site (3' ss) (consensus splice sites; reviewed in (1, 2)). However, these features are not always sufficient to ensure accurate exon definition when exons, which usually average 150nt in length, are separated by intronic regions of up to 500,000nt. Additional sequence elements in introns (eg intronic splice enhancers (ISEs)) and in exons (eg exonic splice enhancers (ESEs)) can act to enhance exon recognition, whilst intronic splice silencers (ISSs) and exonic splice silencers (ESSs), can repress exon recognition (3).



**Figure 1.** Splicing Consensus Sequences and Interactions. a. Location of splice site consensus sequences. Exons are indicated by red boxes, introns by a black line. Invariant nucleotides from consensus sequences are shown in black beneath the introns. b. Interaction between the sequence elements in RNA and corresponding proteins. Grey and black arrows indicate protein-protein interactions promoting splicing, red indicates inhibitory interactions.

Other sequence elements can influence splicing by affecting RNA secondary structure. For example, the weak 5' ss of *SMN2* exon 7 forms part of a stem-loop structure that impedes its recognition by the spliceosome thus repressing splicing (4). Structural elements located deep within introns function to ensure skipping of pseudoexons, and as such, disruption of these sequences can result in abnormal inclusion of pseudoexons in the final transcript. Examples of genes in which this has been reported include *ATM* and *CFTR* (5).

### 2.1.1. Splice site consensus sequences

Splice site consensus sequences were extensively surveyed in vertebrates during the 1980s, and are well conserved (6, 7) (see Figure 1). The splice site consensus sequences can vary in strength, and in the case of weaker signals, other sequence elements may be required for exon recognition. The splice donor site consensus sequence is recognised by U1 snRNP, the branch point is recognised by U2 snRNP, the polypyrimidine tract and splice acceptor site are bound by U2AF.

### 2.1.2. Exonic splicing elements

Exonic splicing enhancers (ESEs) are sequence elements that promote exon inclusion and are required for constitutive splicing and are present in most, if not all, exons (8, 9). Exonic splicing silencers (ESSs) are less well characterised. In contrast to ESEs, they repress exon inclusion and appear to be more prevalent in pseudoexons and exons that are frequently skipped (10). ESEs are generally recognised either by members of the SR protein family, or SR-like proteins (Table 1), and ESSs are often recognised by members of the hnRNP family (11).

The ESE prediction programs ESEfinder (12) and RescueESE (13) have recently become available, and there

have now been a number of studies utilising these programs to identify sequence changes that exert pathogenic effects through disrupting ESE motifs. ESEfinder searches for consensus binding sequences for four SR proteins (SF2/ASF, SC35, SRp40 and SRp55) developed using the SELEX (systematic evolution of ligands by exponential enrichment) procedure (9, 12, 14). Input sequences are given scores according to fit with the loose consensus sequences and motifs above a threshold value are predicted to be recognised by the SR proteins and act as ESEs. ESEfinder is currently limited to predicting ESEs recognised by only four of the family of SR proteins. RescueESE searches input sequences for a set of 238 hexamer sequence motifs found to be significantly enriched in exonic sequences when compared to intronic sequences, and also significantly enriched in exons with weak splice signals compared with exons with strong splice signals. The list of hexamers that RescueESE searches for were generated by a purely bioinformatic approach, and experimental validation has been performed for only 10 of the 238 motifs in an *in vivo* minigene system (15). RescueESE is also limited in that it is unable to predict which splicing factors recognise each hexamer. Other limitations of these programs include their inability to take into account the interconnectivity of splicing to other regulatory processes, such as transcription, capping and polyadenylation (reviewed in (16)), which are very likely to influence the function of exonic splice enhancers.

Larger composite sequence elements that regulate splicing have also been characterised, such as CERES in exon 12 of *CFTR* (17). Splicing enhancer predictor programs are unable to detect such elements, and so detection currently relies on techniques such as *in vivo* selection (18) or antisense microwalk (19), both of which have been used to identify splicing elements in *SMN2* exon 7.

**Table 1.** ESE sequence motifs recognised by SR proteins

SR Protein	Motif	Reference
ASF	CRMSGW <sup>1</sup>	9
	RGAAGAAC	26
	AGGACRRAGC	26
SC35	UGCYGY	8
SRp20	WCWWC	27
	CCUCGUCC	8
SRp40	YRCRKM <sup>1</sup>	9
	TGGGAGCRGTYRGCTCGY	109
SRp55	YYWCWSG*	9
9G8	(GAC) <sub>n</sub>	27

<sup>1</sup>Revised consensus sequences as per (1). R: A/G, S: C/G, M: A/C, W: A/U, Y: C/U, K: G/U

Experimental approaches have also been used to identify functional ESS motifs. For example, using an artificial three exon minigene construct, Wang and colleagues were able to identify 141 decamers containing ESS motifs that induced skipping of the middle exon of the construct (10). An additional minigene construct has also been employed to illustrate the ability of ESSs to influence splice site choice in exons with multiple 5' splice sites or 3' splice sites (20).

## 2.2. Splicing factors

Regulation of splicing is mediated by the functional interactions of the above splicing elements with *trans*-acting regulatory factors, including the macromolecular RNA-protein complex known as the spliceosome and a wide range of RNA binding proteins, including SR proteins and hnRNPs.

### 2.2.1. Spliceosome

The molecular machinery responsible for accurate splicing is the spliceosome and it recognises intron/exon boundaries in the pre-mRNA transcript and is able to catalyse the cut-and-paste reactions of intron removal. Splicing is reliant on the accurate recognition of authentic splice sites from the many pseudo-splice sites present, and the ability of the spliceosome to recognise appropriate splice site consensus sequences relies on the complex coordinated network of sequence signals embedded in both introns and exons surrounding the splice sites and the proteins that interact with them (3, 11, 21) (see Figure 1).

### 2.2.2. SR proteins

SR proteins are highly conserved RNA-binding proteins characterised by one or more amino-proximal RRM-type RNA binding motifs, and a distinctive carboxyl-terminal domain rich in arginine-serine dipeptides (RS domain) of variable length (reviewed in (22, 23)). Members of the SR protein family include: ASF/SF2, SC35, 9G8, SRp30c, SRp20, SRp40, SRp46, SRp55, SRp75, and p54 (reviewed in (23)). The highly conserved RRM domains bind to the pre-mRNA, while the RS domain binds proteins and recruits the splicing apparatus and aids splice site pairing (24). These properties aid in splice site recognition and promote the inclusion of the exon in the final mRNA. SR proteins predominantly bind to exonic splicing enhancers (ESEs; see Section 2.2.2) however in certain circumstances they are also able to bind 'intronic splice enhancers' (ISEs) to promote recognition of nearby splice sites (25).

SR proteins have been shown to influence the selection of alternative splice sites in a concentration dependent manner (9, 14, 26-29). An increased SR protein concentration results in the selection of the proximal 5' ss in pre-mRNAs with multiple 5' ss (30, 31). Tissue specific variations in the total and relative amounts of SR proteins or their mRNAs have been described (28, 32) and SRp55, as one example, is expressed at low levels in the brain which contributes to the tissue specific alternative splicing profile of calcitonin mRNA (33). SR protein expression changes during T-cell activation, and there are differences in concentrations of a number of SR proteins in memory T-cells compared to activated T-cells (34) with concomitant changes in alternatively spliced isoforms of CD44 and CD45 (35). A direct link between these changes in SR protein expression and CD44 and CD45 isoform expression has not yet been established, although hnRNP A1 and hTra2-β regulate splicing of CD44 variable exon 5 (36, 37). Also, there are some tissue specific splicing factors, such as NOVA-1, which is only expressed in neurons (38). It has been proposed that variations in splicing factor expression between tissues is the major contributing factor to tissue-specific alternative splicing (24, 39).

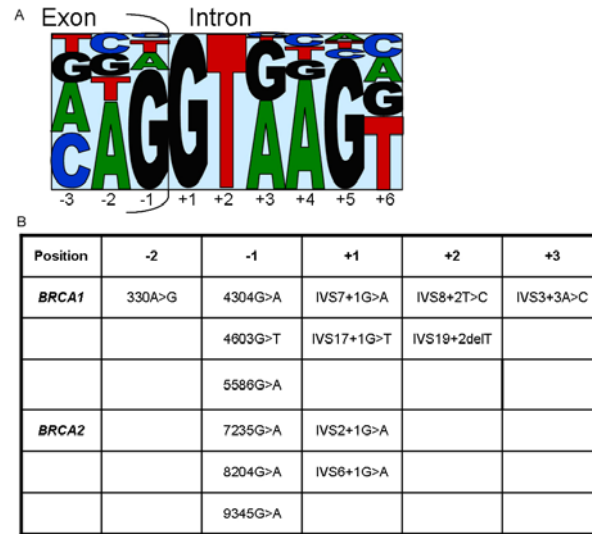
### 2.2.3. hnRNPs

Heterogeneous nuclear ribonucleoproteins (hnRNPs) have been implicated in multiple aspects of RNA processing including splicing, transport, mRNA stability, translational silencing and polyadenylation (reviewed in (40)). hnRNP A1 and hnRNP I (Pyrimidine Tract Binding Protein; PTB) are the most well characterised splicing factors and are associated with exon skipping, although some hnRNPs have been shown to act as splicing activators (41). hnRNPs are thought to block exon recognition by simply preventing SR proteins from binding to their sites in the exonic region, resulting in exon skipping (42). Like SR proteins, the activity of hnRNPs is concentration dependent and also affected by phosphorylation. The concentration of individual hnRNPs also varies across cell and tissue types (43).

Overexpression of hnRNPs has the opposite effect to overexpression of SR proteins on splice site selection (24), suggesting that the silencing ability of hnRNPs is concentration dependent in the same manner as SR proteins and their enhancing ability. It has also been shown that splice site selection in protein 4.1R exon 16 alters with hnRNP A/B protein concentrations during erythropoiesis (44).

## 3. DEFECTS IN SPLICING OF CANCER ASSOCIATED GENES

Defects in splicing have now been implicated in many diseases, such as growth hormone deficiency, Frasier syndrome, Parkinson's disease, cystic fibrosis, retinitis pigmentosa, spinal muscular atrophy, myotonic dystrophy and a large number of cancers (reviewed in (21, 45-47)). Splicing defects can arise as a result of sequence mutations that disrupt splicing elements or through the unbalanced expression or normal alternative splice variants.



**Figure 2.** Mutations in the Splice Donor Sites of *BRCA* genes. a. Schematic representation of the splice site donor consensus sequence. The size of each letter represents the percentage frequency of the nucleotide at each position, as per Figure 1. Numbers indicate the position in the consensus sequence. b. Reported mutations in *BRCA1* and *BRCA2* splice donor sites that alter splice site usage.

### 3.1. Mutations in splicing elements

It has been estimated that around 15% of point mutations that result in disease cause defects in RNA splicing (48). This estimate is only based on classical consensus sequences and does not consider other splicing motifs such as ESEs, so the real proportion may be closer to 50% (1). Most single nucleotide mutations that alter constitutive splicing generate aberrant transcripts as a result of exon skipping or the use of cryptic splice sites. In many cases these result in loss of expression through generation of isoforms that lose the open reading frame and are targeted for nonsense mediated decay (NMD; see Section 4.1).

#### 3.1.1. Mutations in consensus splice sites

Small deletions and single nucleotide mutations that disrupt normal splicing processes by adversely affecting splicing consensus sequences have been detected in a range of cancer-related genes, although not all single nucleotide changes in the splice consensus sequences result in abnormal splicing (for e.g. (49) and (50)).

Some examples of mutations at the 3' end of exons that disrupt splice sites are found in the DNA mismatch repair gene *MLH1*. *MLH1* 667G>A is located in the last nucleotide of exon 8, and 1667G>T is the final base of exon 14. However, two different effects are seen from these mutations. 667G>A results in complete skipping of the 89bp exon 8 which creates a frameshift, and the aberrant transcript is likely to be subjected to NMD resulting in a loss of expression (51). 1667G>T does not result in exon skipping, however the allele generates a transcript retaining 88bp of intron 14 before utilising a cryptic splice site. This intronic region contains an in-frame stop codon, and the aberrant splice transcript is also likely to be subject to NMD (51). In the breast cancer susceptibility gene *BRCA1*, the silent mutation *BRCA1* 4304G>A is located at the last nucleotide of exon 12 and results in complete skipping of the exon leading to a loss of

the ORF (52), and the ORF is also lost by skipping of *BRCA2* exon 13 induced by the mutation 7235G>A, which is located at the final base of the exon (53). A sequence change in the last base of *BRCA1* exon 23, 5586G>A, generates transcripts lacking exon 23 which results in a frameshift and an altered C-terminus of the protein without induction of NMD as exon 24 is the final exon (54). Some exon skipping events arising due to mutation still retain the ORF, for example *BRCA2* 8204G>A is located at the final base of exon 17 and encodes a protein lacking the 57 amino acids encoded by the exon (55). Splice site mutations can also invoke the use of cryptic splice sites, such as *BRCA1* 330A>G, which is located at the second last base of exon 5. This mutation results in the use of a cryptic splice site 22nt upstream in preference to the authentic splice site, leading to a loss of the ORF and a PTC in exon 6 (56).

Mutations in the first base of an exon can also disrupt splicing consensus sequences. The 'missense' mutation 423G>T in the first nucleotide of exon 4 of the tumour suppressor gene *APC* results in skipping of the exon (49).

Intronic mutations in splice donor sequences can also disrupt splicing. A number of examples can be found involving the tumour suppressor *PTEN*. A mutation in the first base of intron 3, IVS3+1G>A, results in complete skipping of exon 3, as does the mutation at position 5 of the intron, IVS3+5G>A (50). The *PTEN-Δ3* transcript maintains the ORF, but has severely reduced protein phosphatase ability (50). The mutation IVS3-1G>A is located in the final nucleotide of *PTEN* intron 3, and leads to complete skipping of exon 4, a loss of the ORF and induction of the NMD pathway (50). Mutations in the splice donor site consensus sequences of *BRCA1* and *BRCA2* have been shown to create a range of splicing abnormalities (Figure 2). Aberrant transcripts generated in this manner that lose the ORF can occur by exon skipping



**Table 2.** Mutations in *NF1* that alter splicing and their predicted effects on ESE motifs

Exon	Nt Change	Amino Acid Change	Position in exon <sup>1</sup>	ESEfinder	RescueESE	% of Exon Skipping (WT) <sup>2</sup>	% of Exon Skipping (Mutant) <sup>2</sup>
7	910C>T	Arg304X	78-30	Increased score of SRp40	Loss of 2 hexamers, creation of 2 hexamers	27.7	88.1
7	943C>T	Gln315X	55-120	Loss of SF2/ASF and SC35	-	27.7	73.2
7	1007G>A	Trp336X	119-56	-	Creation of 1 new hexamer	27.7	79.5
30	5719G>T	Glu1907X	173-31	-	Loss of 5 hexamers	29.4	88.5

Adapted from (67). <sup>1</sup> First number indicates the distance in nt to 3' splice site, and the second number indicates the distance in nt to the 5' splice site. <sup>2</sup> Based on minigene construct splicing assay.

potential to disrupt ESEs and create splicing abnormalities. These mutations resulted in increased exon skipping from 4-100fold above normal levels (67). The exons were then screened using ESEfinder and RescueESE for potential ESE motifs that were disrupted by the mutations. Some of the mutations resulted in clear losses of predicted ESEs, while others yielded more ambiguous results and a selection is shown in Table 2. This study highlights the limitations of the current ESE prediction programs, which arise as a result of the highly degenerate nature of ESE motifs and the inability to factor in other important variables, such as distance from the splice site and splice site strength and effects of simultaneous gene regulatory events.

Only a small number of mutations in the *BRCA* genes have been verified as disrupting ESE motifs leading to splicing abnormalities. A nonsense mutation in exon 18 of *BRCA1*, 5199G>T, has been shown to disrupt an ESE motif recognised by SF2/ASF and only generate transcripts that skip the exon (66). There is a single nucleotide change in *BRCA2* exon 18, 8393C>G, that disrupts three predicted ESE motifs and generates transcripts lacking the exon (68). A number of unclassified sequence variants in *BRCA1* and *BRCA2* have been predicted to disrupt ESE motifs however these remain to be verified experimentally (69).

### 3.1.3. Mutations that create or disrupt other splicing sequences

Single nucleotide changes can alter splicing processes independent of the splicing consensus sequences and ESE motifs. A SNP in *KLF6* which is associated with increased prostate cancer risk (IVS1-27G>A) generates a novel SRp40 binding site in the intron, and increases transcription of three alternatively spliced isoforms which, when translated, are mislocalised to the cytoplasm and antagonise the function of the authentic isoform (70). A mutation deep in intron 2 of *PTEN*, IVS2-38insG, disrupts the consensus branch site sequence, and leads to skipping of exon 3 (50).

Mutations have also been identified in *BRCA1* that alter splicing processes independent of authentic splice sites and ESEs. *BRCA1* 309T>G is predicted to cause a missense change in the highly conserved RING domain of BRCA1, however the mutation promotes the use of a cryptic splice site in exon 5 which is 22bp prior to the authentic 3' ss. Utilisation of this cryptic splice site results in a loss of the reading frame and the transcript is predicted to be targeted by NMD (71). The mutation *BRCA1* IVS5-12G>A creates a new cryptic splice site new site 11nt prior

to exon 6 in *BRCA1*, leading to incorporation of 11nt of the intron and therefore loss of the ORF (72).

### 3.2. Dysregulated expression of splicing factors

Alteration in the concentration of specific splicing factors is believed to contribute to the increased alternative splicing isoforms generated in disease. A number of studies have now shown that the expression levels of individual SR proteins can differ between tumour tissue and the corresponding normal tissues. For example, the expression of SRp20 is induced during development of mammary carcinogenesis, exhibiting a progressive increase in expression as normal cells which alter to pre-neoplasias and then progress to tumours (73). A more recently detected member of the SR protein family, SR-A1, has been found to be overexpressed in aggressive ovarian cancers (74). A more extensive study of ovarian tumours found that SC35 and SF2/ASF show a marked increase in expression compared to normal tissue, SRp20 showed a slight increase while SRp40 and SRp55 expression remained constant (75). These studies suggest that there is specific induction of certain splicing factors rather than a general increase in the expression of all the components of the splicing machinery. This specific induction also extends to SR-like proteins, as hTra2β expression is increased in ovarian tumours, while hTra2α expression remains constant (75).

The alteration of normal splicing factor expression levels is likely to induce changes in splicing across the transcriptome, and already there are numerous reports of increases in expression of alternatively spliced mRNAs from specific genes in tumour tissue compared with normal counterparts. One example is *TSG101*, expression of which is important for genomic stability, cell cycle regulation and suppression of malignant growth. While normal cells express a range of *TSG101* isoforms (76, 77), there have been numerous observations of an increase in the expression of multiple alternatively spliced transcripts in tumour samples from breast, ovary, prostate, thyroid, and cervical tissues (for e.g. (76-79)). Consequences of genes that undergo increased alternative splicing in disease are described in Section 4.2. What remains unclear is whether altered splicing is an early event trigger of tumourigenesis, or whether it is symptomatic of the general breakdown of cellular processes during disease progression. Related to this is the question of what triggers this change from the normal splicing factor concentrations to generate the altered expression of splice variants.

## Pre-mRNA splicing aberrations and cancer

**Table 3.** Cancer-associated alternatively spliced isoforms and functional implications

Gene	Isoform	mRNA and protein coding	Expression data	Functional consequences	Ref
<i>AIB1</i>	AIB1-Δ3	Loss of exon 3 (173bp) results in loss of ORF after 90 amino acids. However, protein synthesis begins from a downstream initiation site in exon in exon 5. The protein lacks the amino-terminal basic helix-loop-helix and a portion of the PAS (Per-Arnt-Sim homology) dimerization domain.	AIB1-Δ3 mRNA is overexpressed in 7/8 breast cancers.	AIB1-Δ3 protein is more active at promoting estrogen receptor mediated transcription, and sensitises cells to growth factor stimulation.	110
<i>WT1</i>	WT1-KTS	Use of a cryptic splice donor site in exon 9, results in skipping of 9bp encoding KTS between zinc fingers 3 and 4.	Most patients with Frasier Syndrome have mutations that create a deficiency of KTS+ isoforms in favour of KTS- isoforms. KTS- expression is increased in breast tumours relative to full length when compared to normal controls.	WT1-KTS- and WTS-KTS+ isoforms have different RNA and DNA binding specificities, and possess consequently transactivate different genes. Frasier Syndrome patients exhibit kidney and gonadal developmental defects.	98, 99, 111
<i>WT1</i>	WT1-Δ5	Loss of exon 5 (51bp) encoding 17 amino acids between proline rich N-terminus and zinc finger domains.	WT1-Δ5 expression is increased in breast tumours relative to full length when compared to normal controls.	The 17 amino acids may exhibit transcriptional repression ability in reporter assay systems, although functional consequences of skipping exon 5 remain unclear.	100, 112
<i>AR</i>	AR-Δ3	Loss of exon 3 (117bp) is predicted to result in a loss of the second zinc finger and impaired DNA binding to androgen response elements.	AR-Δ3 is expressed in breast cancers and cell lines, but was not detected in normal breast tissue.	AR-Δ3 is predicted to have inhibited transcriptional activation ability, leading to reduced activation of growth inhibitory genes in response to androgen.	101
<i>p53</i>	ΔN-p53 (p47)	p47 is generated from p53 transcripts that retain intron 2, and utilise an initiation codon in exon 4 which lack the first 39 amino acids of p53. p47 lacks the first transactivation domain.	Unknown.	p47 is able to dimerise with p53, inhibits p53 mediated transcription, also contributes to cytoplasmic mislocalisation. p47 impairs p53-mediated growth suppression. Also unable to complex with mdm2 and accumulate in response to DNA damage.	113, 114
<i>p73</i>	ΔN-p73	ΔN-p73 is an N-terminal deletion isoform which arises from incorporation of additional exon (exon 3') and use of translational start site in 3'. This isoform lacks the putative transactivation domain of p73.	ΔN-p73 is frequently overexpressed in cancers (ovary, endometrium, cervix, vulva, vagina, breast, kidney, and colon) but not normal tissues.	ΔN-p73 is able to suppresses p53 and p73 transcriptional activation. ΔN-p73 is also able to inhibit p53 suppression of colony formation.	115-117
<i>WISP1</i>	WISP1v	Lacks exon 3 which results in loss of a module of Von Willebrand type C which is thought to participate in protein complex formation.	WISP1v was found to be overexpressed in 18 of 21 scirrhous gastric carcinomas.	Overexpression results in increased cellular transformation and rapid growth. Overexpression of WISP1v also enhanced the invasive phenotype of cultured gastric carcinoma cells, allowing invasion of cells through collagen <i>in vitro</i> whereas the full length WISP1 protein did not induce invasion.	118
<i>Osteopontin</i>	Osteopontin-c	Lacks exon 4 (174bp), resulting in a loss of 58 amino acids.	Osteopontin-c is not normally expressed in mammary tissue, however is found expressed in breast tumours.	Overexpression of osteopontin-c enhances clone formation in soft agar, and also a loss of adhesion compared to overexpression of the full length osteopontin.	119
<i>mdm2</i>	mdm2-Δ4-11	Skipping of exons 4-11 results in a loss of 90% of the p53 interaction domain, the NLS and acidic domain.	Expression of mdm2 splice variants lacking p53 interaction domain has been observed in high grade late stage ovarian and bladder cancers.	Loses ability to bind p53, overexpression shows transforming ability.	120-122

Ref: Reference

### 3.3. Dysregulated expression of splice variants

Changes in the endogenous alternative splicing patterns of genes in cancer may contribute to disease progression via a range of mechanisms that confer a selective advantage to cells housing such changes. Proteins generated by alternatively spliced mRNA isoforms may act as dominant negatives on their full length counterparts, while others may have entirely novel functional properties. Both of these eventualities can have dramatic implications for disease progression. Table 3 summarises some examples of splice variants overexpressed in cancer that possess functional properties that distinguish them from their full

length counterparts and implicate the isoforms in disease progression, whilst the text below highlights some specific examples.

The breast cancer susceptibility gene *BRCA1* consists of 24 exons and expresses a large number of splice variants in a tissue-specific manner (reviewed in (80)). Several studies have found that four predominant *BRCA1* mRNA variants—full length, Δ9-10, Δ11q, and Δ9-10-11q—are expressed in a variety of tissues under different conditions (81-83). Another significant splice variant of *BRCA1* is *BRCA1-Δ11*, which is also expressed in the



## Pre-mRNA splicing aberrations and cancer

mouse (84). These splice variants all maintain the open reading frame, and their protein products retain some of the full length activity in functional assays.

The expression of alternatively spliced *BRCA1* isoforms in normal and malignant tissue is poorly understood. Conflicting studies have confounded attempts to characterise *BRCA1* splicing in such a context. For example, in normal and malignant breast epithelial cells, the full length *BRCA1* and  $\Delta 11q$  isoforms were the higher expressed isoforms, compared to  $\Delta 9-10$  and  $\Delta 9-10-11q$  (83), whereas in other breast and ovarian tumours the  $\Delta 9-10$  variant was highly expressed and lower amounts of  $\Delta 11q$  were detected (85). The conflicting results of these studies may reflect the heterogeneity of different tumour samples as well as different methodologies. The expression of the predominant *BRCA1* splice variants has also been examined across multiple cell lines using a multiplex PCR protocol (81). Expression of the full length *BRCA1* made up ~80% of the pool of *BRCA1* transcripts in normal breast tissue, whereas in MCF7 cells, a tumourigenic mammary epithelial cell line, expression was reduced to 50% of the total pool. In MDA-MB-231 cells, a more aggressive tumourigenic mammary epithelial cell line expression of the full length was reduced to ~35%. With the decrease of full length *BRCA1* expression, the expression of *BRCA1* $\Delta 9-10$  increased from 8% in normal breast tissue to 24% in MCF7s, and 44% in MDA-MB-231s. This increase in expression correlates to increased invasive activity of the cell lines (86), or it may be representative of cell type as MCF7 cells display a luminal phenotype whereas MDA-MB-231 cells express a range of mesenchymal markers.

*BRCA1-IRIS* is another *BRCA1* splice variant and is generated by continuation of the transcript into intron 11 in preference to splicing exon 11 to exon 12. The transcript encodes a protein which has an identical 1365 amino acids to the full length *BRCA1*, prior to a novel 34 amino acid C-terminus. Expression of the two isoforms is inconsistent across both the cell cycle and tissue types, with *BRCA1-IRIS* mRNA expression five times higher in white blood cells than the full length (87). In response to acute DNA damage full length *BRCA1* forms nuclear foci, while *BRCA1-IRIS* does not, and only the full length *BRCA1* protein associates with *BARD1* (87). *BRCA1-IRIS* is exclusively chromatin-associated and is a positive influence on DNA replication, as it associates with proteins that bind to initiation sites of DNA replication, such as *ORC1*, *Cdc6* and *MCM2* (87). In a follow up study, Nakuci and colleagues were able to show that *BRCA1-IRIS* also forms complexes with *SRC1* and *SRC3*, nuclear receptor co-activators, and is recruited to the cyclin D1 promoter, inducing cyclin D1 expression and in turn, cellular proliferation (88). This is in contrast to a number of other studies showing that full length *BRCA1* can inhibit cell cycle progression and in turn cellular proliferation (89-91). These data suggest that despite the sequence similarity, *BRCA1-IRIS* may act as a proto-oncogene while the full length *BRCA1* acts as a tumour suppressor.

The breast cancer susceptibility gene 2 (*BRCA2*) consists of 27 exons and to date only nine isoforms,

including the full length, have been reported as expressed in normal tissues. One of these is *BRCA2*  $\Delta 3$ , which has been shown to be expressed in a range of normal tissues as well as in cancerous samples (92). While the skipping of *BRCA2* exon 3 maintains the ORF, it results in loss of the transcriptional activation domain (92, 93). The protein domain encoded by *BRCA2* exon 3 shows sequence similarity to the activation domain of c-Jun, and this region has been shown to possess transcriptional activation activity (94).

Another *BRCA2* splice variant that has been detected as expressed in normal breast tissue is the *BRCA2- $\Delta 12$*  isoform (95). Exon 12 is 96bp long, and skipping of the exon maintains the open reading frame, however it is unclear if any protein interaction domains are interrupted by this exon skipping event. *Brca2- $\Delta 12$*  is also expressed in the mouse, and this evolutionary conservation suggests that the isoform may perform specific functions in the cell (95). When breast tumour samples were analysed for *BRCA2- $\Delta 12$*  expression, 4/12 showed higher expression of the isoform than their matched normal controls and this higher expression was associated with steroid receptor negativity (95). Further work is required to determine what functional role *BRCA2- $\Delta 12$*  performs and to establish whether there is any correlation between *BRCA2- $\Delta 12$*  overexpression and breast or ovarian cancer.

*CHK2* is a DNA damage checkpoint kinase. *CHK2* is subject to extensive alternative splicing in cancers, with more than 90 variants detected in a panel of 53 breast carcinomas in addition to the full length isoform (96). In normal breast tissue an average of four alternatively spliced isoforms were detected, while in breast tumours an average of six isoforms were found (96). Over half of the splice variants detected contained PTCs, which may make them targets of the NMD pathway, while other splice variants were aberrantly localised to the cytoplasm (96). These two mechanisms, NMD and aberrant subcellular localisation, would significantly reduce the available pool of functioning *CHK2* in these tumours.

*PTEN* is a dual-phosphatase tumour suppressor which is able to dephosphorylate *PIP3*, the *PDK1* activator. Inactivated *PTEN* results in unregulated activation of *Akt/PKB*, resulting in uncontrolled proliferation. *PTEN* also inhibits the *MAPK* pathway, and is able to elicit apoptosis and cell cycle arrest at G1. Inactivation of *PTEN* is found in 85% of Cowden syndrome (characterised by increased risk of breast and thyroid cancer) cases, and *PTEN* inactivation is also associated with a variety of sporadic cancers. There are 11 natural splice variants of *PTEN* including the full length isoform. Eight isoforms were identified by Agrawal and Eng (97). These isoforms have premature termination codons and are expressed normally at varying levels in different tissues. The presence of these PTCs may target these isoforms for NMD, as alternatively spliced isoforms of *PTEN* containing PTCs are rapidly degraded prior to generating any protein (50). Full length *PTEN* expression was shown to be reduced in 10/12 sporadic breast cancers, as were some of the splice variants, while expression of the *PTEN-5b* splice variant



## Pre-mRNA splicing aberrations and cancer

**Table 4.** Alternatively spliced isoforms with biomarker potential

Gene	Isoform	Detail	Expression data	Ref
<i>NRSF</i>	sNRSF	Incorporation of a 50bp pseudoexon between exons 5 and 6, contains a stop codon.	sNRSF is overexpressed in small cell lung cancers, but not expressed in other lung cancers.	123
<i>Ron</i>	Ron-Δ11 (ΔRon)	Skipping of Ron exon 11 (147bp) results in an in-frame deletion of 49 amino acids, the isoform is constitutively active and does not undergo proteolytic cleavage.	ΔRon mRNA accumulates in breast and colon cancer cells.	124
<i>CD44</i>	CD44-intron9	Retention of intron 9.	Intron 9 retention only occurs in ovarian cancer tissue.	125

Ref: Reference

mRNA was increased in 9/12 sporadic breast cancers (97). The increased alternative splicing of PTEN results in the generation of a number of isoforms containing PTCs and these are predicted to be targets of the NMD pathway, which may be an effective means of severely reducing the amount of functional PTEN in the tumour cells.

WT1-KTS- is an alternatively spliced isoform of WT1 which is overexpressed relative to the WT1-KTS+ isoform in patients with Frasier Syndrome. In the majority of cases, this imbalance of alternative splicing arises due to mutations in WT1 which disrupt normal KTS+ splicing (see for e.g. (98)). These two isoforms of WT1 possess transactivation activity, however transactivate a different selection of target genes (99). This is an example of splicing disruption arising from mutations in the gene which leads to disease. Other cases of disease-associated disruptions in alternative splicing do not necessarily arise due to gene mutations, with WT1-Δ5, another WT1 isoform overexpressed in breast tumours relative to normal tissues (100). As WT1-Δ5 is expressed normally, this is an example of disrupted regulation of alternative splicing, leading to an increase in expression of one particular isoform. The WT1-Δ5 is hypothesized to have a reduced capacity to repress transcriptional activation compared to the full length isoform (100). In some cases, certain alternatively spliced isoforms are found to be expressed in tumours originating from tissues that do not express the alternatively spliced isoform normally. Androgen receptor (AR) has been shown to express an isoform lacking exon 3 in breast cancers and breast cancer cell lines, however no isoform is detectable in normal breast tissue (101). The AR-Δ3 protein product is predicted to confer a reduced activation of growth inhibitory genes compared to the full length AR (101).

### 4. CLINICAL IMPLICATIONS OF SPLICING DEFECTS IN CANCER

#### 4.1. Tumour biomarkers

The expression of alternatively spliced isoforms that are unique to cancerous tissue can present biomarkers for disease and also treatment targets. Detecting the expression of these unique isoforms may have clear diagnostic and prognostic value. In addition, these novel isoforms may supply cancer specific epitopes able to be recognised by the immune system, and may serve as targets for immunotherapy. Over 1450 different potential tumour antigens have been identified using the SEREX approach (serological identification of antigens by recombinant expression cloning) (reviewed in (47)). Some examples of alternatively spliced isoforms that may potentially act as biomarkers are shown in Table 4. Other potential

biomarkers are listed in Table 3, such as AR-Δ3 and osteopontin-c, both potential biomarkers of breast cancer.

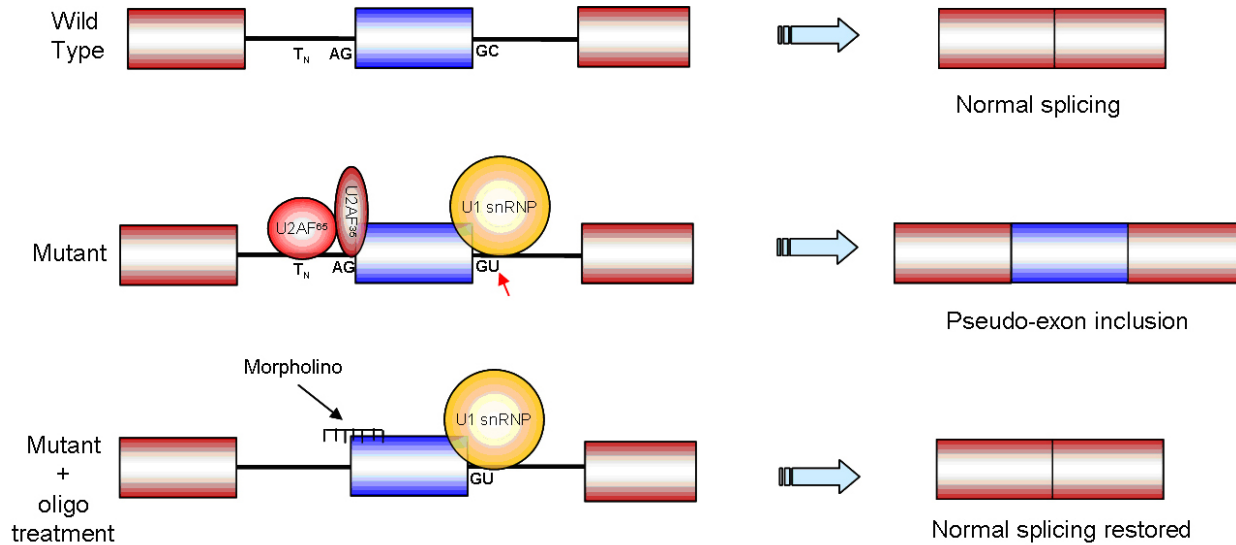
#### 4.2. Therapeutic strategies for correcting splicing defects

New techniques are currently being established to correct splicing abnormalities that are associated with disease. One such technique involves the use of oligonucleotides targeted to splicing sequence motifs. These oligonucleotides, commonly morpholinos, are resistant to nucleases and do not allow degradation of mRNA in RNA-oligonucleotide hybrids by RNase H. The clinical implications of correcting splicing abnormalities are illustrated in the following examples.

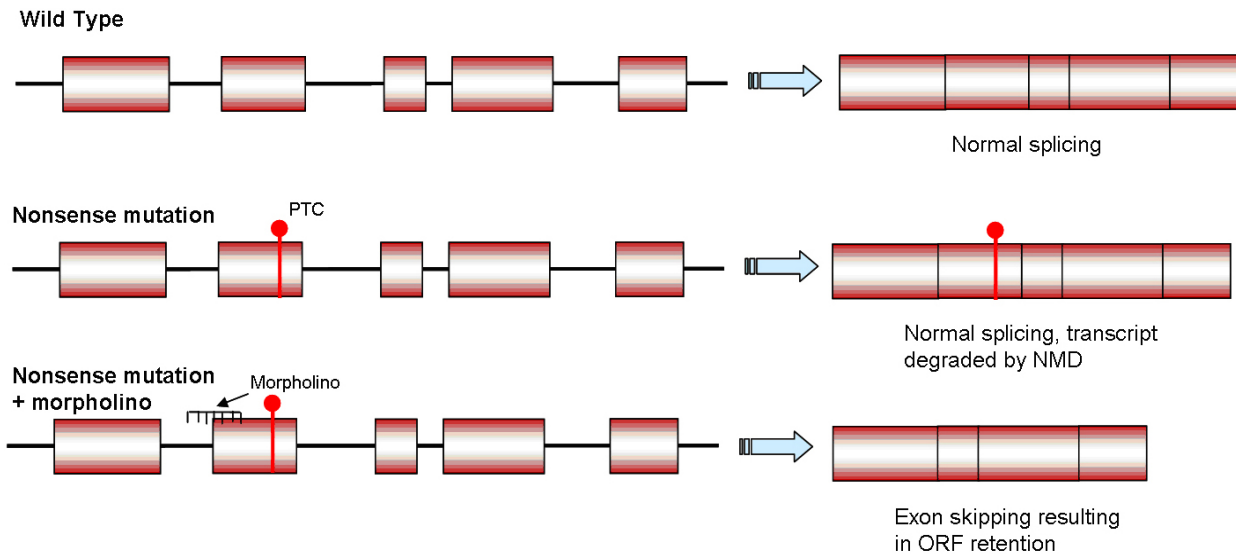
β-thalassemia results from mutations that cause defective β-globin expression, and result in protein deficiency. Splicing mutations are among the most common mutations causing defective β-globin expression, and some of the most common splicing mutations activate aberrant splice sites, even though correct sites remain potentially functional. Two mutations located deep within intron 2, IVS2-654C>T and IVS2-745C>G, both create an aberrant 5' splice site and use a common cryptic 3' splice site at 579, leading to inclusion of intronic sequence in the β-globin mRNA which includes a stop codon preventing proper translation of the mRNA. The ultimate consequence of these mutations is β-thalassemia resulting from β-globin deficiency. Antisense morpholinos to these aberrant splice sites in mutant β-globin pre-mRNAs were able to restore correct splicing and translation of β-globin mRNA in mononuclear cells from peripheral blood of thalassemic patients (102) (see Figure 4). The repaired expression reached around 30% of that of the normal, and was still detectable nine days after morpholino treatment.

Antisense oligonucleotides have also been used to re-direct dystrophin pre-mRNA processing by blocking sequences crucial to splicing. Duchenne and Becker muscular dystrophies arise due to mutations in dystrophin, with Duchenne muscular dystrophy usually characterised by the absence of functional dystrophin protein. Morpholinos targeted to the dystrophin pre-mRNA are able to induce skipping of specific exons of the pre-mRNA transcript in order to maintain the ORF, which may be disrupted by frameshift mutations or splicing mutations, or to skip nonsense mutations (103-105) (see Figure 5). These morpholinos can target one exon specifically, or are able to be used in a cocktail to induce skipping of multiple exons. Morpholinos can be targeted to the splice consensus sequences, however they are marginally more effective when targeted to ESE motifs (103). Morpholinos are also

## Pre-mRNA splicing aberrations and cancer



**Figure 4.** Restoration of  $\beta$ -globin splicing in mutation carriers. The mutation  $C>T$  is marked by a red arrow, creating a new splice donor consensus sequence, and inclusion of the pseudoexon. Addition of a morpholino antisense to the splice consensus sequence of the pseudoexon blocks binding of snRNPs and results in skipping of the pseudoexon (102).



**Figure 5.** Restoration of the dystrophin ORF by inducing exon skipping. Exons containing PTCs are targeted by morpholinos antisense to consensus splice sites (or ESEs) which prevent splice site recognition and result in exon skipping. The ORF is restored and protein production is facilitated.

able to be injected into muscular tissue in the absence of a delivery reagent (103).

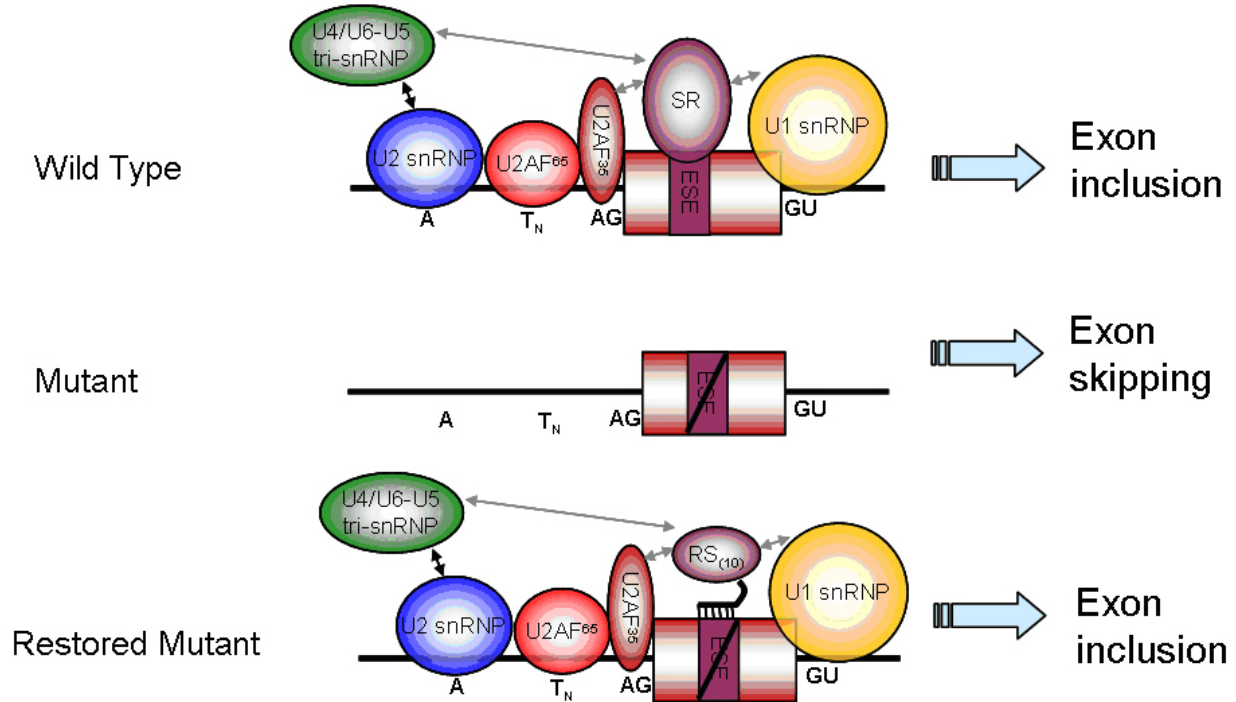
While morpholinos are able to induce skipping of exons, enhancing the inclusion of exons has also been examined. Cartegni and Krainer were able to restore splicing in two systems involving ESEs which had been abrogated by single nucleotide sequence changes (106) using a PNA-peptide approach termed 'ESSENCE' (exon-specific splicing enhancement by small chimeric effectors). A peptide nucleic acid (PNA) sequence complementary to

the mutant pre-mRNA fused to a peptide containing 10 RS repeats was able to act in place of an SR protein and restore splicing in both a *BRCA1* minigene system and an *SMN2* minigene system (see Figure 6). This study shows that specifically designed molecules may be able to restore splicing of exons that would otherwise be skipped.

## 5. CONCLUSIONS AND PERSPECTIVES

The dramatic effects of dysregulated splicing can arise due to *cis* sequence mutations in the pre-mRNA, or an

## Pre-mRNA splicing aberrations and cancer



**Figure 6.** Restoration of exon inclusion in instances of mutated ESEs. In the restored mutant the PNA oligo fused to the RS repeats binds to the mutant sequence and is able to carry out the protein-protein interactions of the normal SR protein (106).

imbalance in *trans* splicing factors or splice variants. Future research in the field of splicing and cancer will include more reliable identification of functional splicing elements and a better understanding of the extent and functional consequences of alternate splicing in cancer-associated genes and the contribution such variants make to the tumourigenic process.

In terms of identifying functional splice elements, there is much interest in the role of ESEs. It has been suggested that ESEs are present in all exons, whether constitutively or alternatively spliced (1). However, the computer programs that predict ESE motifs often yield high numbers of false positive results and ambiguous results (67, 69). The ability to identify functional ESE motifs from these outputs results remains a challenge, however applying filters to the output data of these predictions programs may improve their accuracy (69).

Alternative isoforms can have pathogenic effects through either deficiency in the amount of functional full length protein or through novel characteristics or altered functional capabilities to those of the full length. A complete characterization of the expression and function of all splice variants arising from all cancer associated genes will be required in order to determine the precise role of splice variants in cancer. Projects such as the ENCODE project will achieve this (107). Data from the recently described pilot ENCODE certainly underscore the number and diversity of alternate transcripts, predicting an average of 5.4 alternate transcripts per locus (108). Future priorities include determining the functions of these alternate

transcripts and their relative expression in normal and tumour cells.

## 6. ACKNOWLEDGEMENTS

This work was supported by funds from The University of Queensland and The National Health and Medical Research Council of Australia.

## 7. REFERENCES

1. L. Cartegni, S. L. Chew, A. R. Krainer. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet*, 3(4), 285-298 (2002)
2. M. L. Hastings, A. R. Krainer. Pre-mRNA splicing in the new millennium. *Curr Opin Cell Biol*, 13(3), 302-309 (2001)
3. R. H. Hovhannisyanyan, R. P. Carstens. A novel intronic cis element, ISE/ISS-3, regulates rat fibroblast growth factor receptor 2 splicing through activation of an upstream exon and repression of a downstream exon containing a noncanonical branch point sequence. *Mol Cell Biol*, 25(1), 250-263 (2005)
4. N. N. Singh, R. N. Singh, E. J. Androphy. Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Res*, 35(2), 371-389 (2007)
5. E. Buratti, A. Dhir, M. A. Lewandowska, F. E. Baralle. RNA structure is a key regulatory element in pathological ATM and CFTR pseudoexon inclusion events. *Nucleic Acids Res* (2007)

6. M. B. Shapiro, P. Senapathy. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res*, 15(17), 7155-7174 (1987)
7. R. A. Padgett, M. M. Konarska, M. Aebi, H. Hornig, C. Weissmann, P. A. Sharp. Nonconsensus branch-site sequences in the *in vitro* splicing of transcripts of mutant rabbit beta-globin genes. *Proc Natl Acad Sci U S A*, 82(24), 8349-8353 (1985)
8. T. D. Schaal, T. Maniatis. Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. *Mol Cell Biol*, 19(3), 1705-1719 (1999)
9. H. X. Liu, M. Zhang, A. R. Krainer. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev*, 12(13), 1998-2012 (1998)
10. Z. Wang, M. E. Rolish, G. Yeo, V. Tung, M. Mawson, C. B. Burge. Systematic identification and analysis of exonic splicing silencers. *Cell*, 119(6), 831-845 (2004)
11. C. W. Smith, J. Valcarcel. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci*, 25(8), 381-388 (2000)
12. L. Cartegni, J. Wang, Z. Zhu, M. Q. Zhang, A. R. Krainer. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res*, 31(13), 3568-3571 (2003)
13. W. G. Fairbrother, G. W. Yeo, R. Yeh, P. Goldstein, M. Mawson, P. A. Sharp, C. B. Burge. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res*, 32(Web Server Issue), W187-W190 (2004)
14. H. X. Liu, S. L. Chew, L. Cartegni, M. Q. Zhang, A. R. Krainer. Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. *Mol Cell Biol*, 20(3), 1063-1071 (2000)
15. W. G. Fairbrother, R. F. Yeh, P. A. Sharp, C. B. Burge. Predictive identification of exonic splicing enhancers in human genes. *Science*, 297(5583), 1007-1013 (2002)
16. N. J. Proudfoot, A. Furger, M. J. Dye. Integrating mRNA processing with transcription. *Cell*, 108(4), 501-512 (2002)
17. F. Pagani, C. Stuardi, M. Tzetis, E. Kanavakis, A. Efthymiadou, S. Doudounakis, T. Casals, F. E. Baralle. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum Mol Genet*, 12(10), 1111-1120 (2003)
18. N. N. Singh, E. J. Androphy, R. N. Singh. *In vivo* selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes. *Rna*, 10(8), 1291-1305 (2004)
19. Y. Hua, T. A. Vickers, B. F. Baker, C. F. Bennett, A. R. Krainer. Enhancement of SMN2 Exon 7 Inclusion by Antisense Oligonucleotides Targeting the Exon. *PLoS Biol*, 5(4), e73 (2007)
20. Z. Wang, X. Xiao, E. Van Nostrand, C. B. Burge. General and specific functions of exonic splicing silencers in splicing control. *Mol Cell*, 23(1), 61-70 (2006)
21. N. A. Faustino, T. A. Cooper. Pre-mRNA splicing and human disease. *Genes Dev*, 17(4), 419-437 (2003)
22. A. J. Lopez. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu Rev Genet*, 32(279-305 (1998)
23. B. R. Graveley. Sorting out the complexity of SR protein functions. *Rna*, 6(9), 1197-1211 (2000)
24. J. F. Caceres, A. R. Kornblihtt. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet*, 18(4), 186-193 (2002)
25. M. E. Gallego, R. Gattoni, J. Stevenin, J. Marie, A. Expert-Bezancon. The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the beta-tropomyosin alternative exon 6A. *Embo J*, 16(7), 1772-1784 (1997)
26. R. Tacke, J. L. Manley. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *Embo J*, 14(14), 3540-3551 (1995)
27. Y. Cavaloc, C. F. Bourgeois, L. Kister, J. Stevenin. The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. *Rna*, 5(3), 468-483 (1999)
28. A. M. Zahler, K. M. Neugebauer, W. S. Lane, M. B. Roth. Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science*, 260(5105), 219-222 (1993)
29. J. Wang, J. L. Manley. Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing *in vivo* in diverse ways. *Rna*, 1(3), 335-346 (1995)
30. M. Molin, G. Akusjarvi. Overexpression of essential splicing factor ASF/SF2 blocks the temporal shift in adenovirus pre-mRNA splicing and reduces virus progeny formation. *J Virol*, 74(19), 9002-9009 (2000)
31. J. F. Caceres, S. Stamm, D. M. Helfman, A. R. Krainer. Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. *Science*, 265(5179), 1706-1709 (1994)
32. A. Hanamura, J. F. Caceres, A. Mayeda, B. R. Franza, Jr., A. R. Krainer. Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. *Rna*, 4(4), 430-444 (1998)
33. Q. Tran, T. P. Coleman, J. R. Roesser. Human transformer 2beta and SRp55 interact with a calcitonin-specific splice enhancer. *Biochim Biophys Acta*, 1625(2), 141-152 (2003)
34. R. Lemaire, A. Winne, M. Sarkissian, R. Lafyatis. SF2 and SRp55 regulation of CD45 exon 4 skipping during T cell activation. *Eur J Immunol*, 29(3), 823-837 (1999)
35. G. R. Srean, J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G. Jackson, J. I. Bell, A. R. Krainer. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *Embo J*, 14(17), 4336-4349 (1995)
36. D. O. Watermann, Y. Tang, A. Zur Hausen, M. Jager, S. Stamm, E. Stickeler. Splicing factor Tra2-beta1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene. *Cancer Res*, 66(9), 4774-4780 (2006)
37. N. Matter, M. Marx, S. Weg-Remers, H. Ponta, P. Herrlich, H. Konig. Heterogeneous ribonucleoprotein A1 is part of an exon-specific splice-silencing complex controlled by oncogenic signaling pathways. *J Biol Chem*, 275(45), 35353-35360 (2000)

## Pre-mRNA splicing aberrations and cancer

38. A. D. Polydorides, H. J. Okano, Y. Y. Yang, G. Stefani, R. B. Darnell. A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc Natl Acad Sci U S A*, 97(12), 6350-6355 (2000)
39. A. J. Pollard, A. R. Krainer, S. C. Robson, G. N. Europe-Finner. Alternative splicing of the adenylyl cyclase stimulatory G-protein G alpha(s) is regulated by SF2/ASF and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and involves the use of an unusual TG 3'-splice Site. *J Biol Chem*, 277(18), 15241-15251 (2002)
40. A. M. Krecic, M. S. Swanson. hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol*, 11(3), 363-371 (1999)
41. H. Sun, L. A. Chasin. Multiple splicing defects in an intronic false exon. *Mol Cell Biol*, 20(17), 6414-6425 (2000)
42. J. Zhu, A. Mayeda, A. R. Krainer. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell*, 8(6), 1351-1361 (2001)
43. H. Kamma, D. S. Portman, G. Dreyfuss. Cell type-specific expression of hnRNP proteins. *Exp Cell Res*, 221(1), 187-196 (1995)
44. V. C. Hou, R. Lersch, S. L. Gee, J. L. Ponthier, A. J. Lo, M. Wu, C. W. Turck, M. Koury, A. R. Krainer, A. Mayeda, J. G. Conboy. Decrease in hnRNP A/B expression during erythropoiesis mediates a pre-mRNA splicing switch. *Embo J*, 21(22), 6195-6204 (2002)
45. M. A. Garcia-Blanco, A. P. Baraniak, E. L. Lasda. Alternative splicing in disease and therapy. *Nat Biotechnol*, 22(5), 535-546 (2004)
46. J. P. Venables. Aberrant and alternative splicing in cancer. *Cancer Res*, 64(21), 7647-7654 (2004)
47. Z. Kalnina, P. Zayakin, K. Silina, A. Line. Alterations of pre-mRNA splicing in cancer. *Genes Chromosomes Cancer* (2005)
48. M. Krawczak, J. Reiss, D. N. Cooper. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet*, 90(1-2), 41-54 (1992)
49. S. Aretz, S. Uhlhaas, Y. Sun, C. Pagenstecher, E. Mangold, R. Caspari, G. Moslein, K. Schulmann, P. Propping, W. Friedl. Familial adenomatous polyposis: aberrant splicing due to missense or silent mutations in the APC gene. *Hum Mutat*, 24(5), 370-380 (2004)
50. S. Agrawal, R. Pilarski, C. Eng. Different splicing defects lead to differential effects downstream of the lipid and protein phosphatase activities of PTEN. *Hum Mol Genet*, 14(16), 2459-2468 (2005)
51. A. Sharp, G. Pichert, A. Lucassen, D. Eccles. RNA analysis reveals splicing mutations and loss of expression defects in MLH1 and BRCA1. *Hum Mutat*, 24(3), 272 (2004)
52. K. Claes, B. Poppe, E. Machackova, I. Coene, L. Foretova, A. De Paepe, L. Messiaen. Differentiating pathogenic mutations from polymorphic alterations in the splice sites of BRCA1 and BRCA2. *Genes Chromosomes Cancer*, 37(3), 314-320 (2003)
53. M. Thomassen, T. A. Kruse, P. K. Jensen, A. M. Gerdes. A missense mutation in exon 13 in BRCA2, c.7235G>A, results in skipping of exon 13. *Genet Test*, 10(2), 116-120 (2006)
54. A. Ladopoulou, I. Konstantopoulou, S. Armaou, E. Efstathiou, M. Mihalatos, G. Nasioulas, G. Bardi, N. Pandis, D. Yannoukakos, Athens. A change in the last base of BRCA1 exon 23, 5586G-->A, results in abnormal RNA splicing. *Cancer Genet Cytogenet*, 134(2), 175-177 (2002)
55. W. Hofmann, D. Horn, C. Huttner, E. Classen, S. Scherneck. The BRCA2 variant 8204G>A is a splicing mutation and results in an in frame deletion of the gene. *J Med Genet*, 40(3), e23 (2003)
56. A. Vega, B. Campos, B. Bressac-De-Paillerets, P. M. Bond, N. Janin, F. S. Douglas, M. Domenech, M. Baena, C. Pericay, C. Alonso, A. Carracedo, M. Baiget, O. Diez. The R71G BRCA1 is a founder Spanish mutation and leads to aberrant splicing of the transcript. *Hum Mutat*, 17(6), 520-521 (2001)
57. M. T. Pyne, D. Pruss, B. E. Ward, T. Scholl. A characterization of genetic variants in BRCA1 intron 8 identifies a mutation and a polymorphism. *Mutat Res*, 406(2-4), 101-107 (1999)
58. F. Bonatti, C. Pepe, M. Tancredi, G. Lombardi, P. Aretini, E. Sensi, E. Falaschi, G. Cipollini, G. Bevilacqua, M. A. Caligo. RNA-based analysis of BRCA1 and BRCA2 gene alterations. *Cancer Genet Cytogenet*, 170(2), 93-101 (2006)
59. X. Chen, T. T. Truong, J. Weaver, B. A. Bove, K. Cattie, B. A. Armstrong, M. B. Daly, A. K. Godwin. Intronic alterations in BRCA1 and BRCA2: effect on mRNA splicing fidelity and expression. *Hum Mutat*, 27(5), 427-435 (2006)
60. M. S. Brose, P. Volpe, K. Paul, J. E. Stopfer, T. A. Colligon, K. A. Calzone, B. L. Weber. Characterization of two novel BRCA1 germ-line mutations involving splice donor sites. *Genet Test*, 8(2), 133-138 (2004)
61. A. A. Tesoriero, E. M. Wong, M. A. Jenkins, J. L. Hopper, M. A. Brown, G. Chenevix-Trench, A. B. Spurdle, M. C. Southey. Molecular characterization and cancer risk associated with BRCA1 and BRCA2 splice site variants identified in multiple-case breast cancer families. *Hum Mutat*, 26(5), 495 (2005)
62. J. C. Keaton, D. R. Nielsen, B. C. Hendrickson, M. T. Pyne, L. Scheuer, B. E. Ward, A. R. Brothman, T. Scholl. A biochemical analysis demonstrates that the BRCA1 intronic variant IVS10-2A--> C is a mutation. *J Hum Genet*, 48(8), 399-403 (2003)
63. K. Claes, J. Vandesompele, B. Poppe, K. Dahan, I. Coene, A. De Paepe, L. Messiaen. Pathological splice mutations outside the invariant AG/GT splice sites of BRCA1 exon 5 increase alternative transcript levels in the 5' end of the BRCA1 gene. *Oncogene*, 21(26), 4171-4175 (2002)
64. T. Scholl, M. T. Pyne, D. Russo, B. E. Ward. BRCA1 IVS16+6T-->C is a deleterious mutation that creates an aberrant transcript by activating a cryptic splice donor site. *Am J Med Genet*, 85(2), 113-116 (1999)
65. K. Nagao, N. Togawa, K. Fujii, H. Uchikawa, Y. Kohno, M. Yamada, T. Miyashita. Detecting tissue-specific alternative splicing and disease-associated aberrant splicing of the PTCH gene with exon junction microarrays. *Hum Mol Genet*, 14(22), 3379-3388 (2005)

66. H. X. Liu, L. Cartegni, M. Q. Zhang, A. R. Krainer. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat Genet*, 27(1), 55-58 (2001)
67. A. Zatkova, L. Messiaen, I. Vandenbroucke, R. Wieser, C. Fonatsch, A. R. Krainer, K. Wimmer. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Hum Mutat*, 24(6), 491-501 (2004)
68. J. D. Fackenthal, L. Cartegni, A. R. Krainer, O. I. Olopade. BRCA2 T2722R is a deleterious allele that causes exon skipping. *Am J Hum Genet*, 71(3), 625-631 (2002)
69. C. Pettigrew, N. Wayne, P. K. Lovelock, S. V. Tavtigian, G. Chenevix-Trench, A. B. Spurdle, M. A. Brown. Evolutionary conservation analysis increases the colocalization of predicted exonic splicing enhancers in the BRCA1 gene with missense sequence changes and in-frame deletions, but not polymorphisms. *Breast Cancer Res*, 7(6), R929-R939 (2005)
70. G. Narla, A. Difeo, H. L. Reeves, D. J. Schaid, J. Hirshfeld, E. Hod, A. Katz, W. B. Isaacs, S. Hebring, A. Komiya, S. K. McDonnell, K. E. Wiley, S. J. Jacobsen, S. D. Isaacs, P. C. Walsh, S. L. Zheng, B. L. Chang, D. M. Friedrichsen, J. L. Stanford, E. A. Ostrander, A. M. Chinnaiyan, M. A. Rubin, J. Xu, S. N. Thibodeau, S. L. Friedman, J. A. Martignetti. A germline DNA polymorphism enhances alternative splicing of the KLF6 tumor suppressor gene and is associated with increased prostate cancer risk. *Cancer Res*, 65(4), 1213-1222 (2005)
71. Y. Yang, S. Swaminathan, B. K. Martin, S. K. Sharan. Aberrant splicing induced by missense mutations in BRCA1: clues from a humanized mouse model. *Hum Mol Genet*, 12(17), 2121-2131 (2003)
72. J. D. Hoffman, S. E. Hallam, V. L. Venne, E. Lyon, K. Ward. Implications of a novel cryptic splice site in the BRCA1 gene. *Am J Med Genet*, 80(2), 140-144 (1998)
73. E. Stickeler, F. Kittrell, D. Medina, S. M. Berget. Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. *Oncogene*, 18(24), 3574-3582 (1999)
74. A. Scorilas, L. Kyriakopoulou, D. Katsaros, E. P. Diamandis. Cloning of a gene (SR-A1), encoding for a new member of the human Ser/Arg-rich family of pre-mRNA splicing factors: overexpression in aggressive ovarian cancer. *Br J Cancer*, 85(2), 190-198 (2001)
75. D. C. Fischer, K. Noack, I. B. Runnebaum, D. O. Watermann, D. G. Kieback, S. Stamm, E. Stickeler. Expression of splicing factors in human ovarian cancer. *Oncol Rep*, 11(5), 1085-1090 (2004)
76. M. P. Lee, A. P. Feinberg. Aberrant splicing but not mutations of TSG101 in human breast cancer. *Cancer Res*, 57(15), 3131-3134 (1997)
77. S. A. Gayther, P. Barski, S. J. Batley, L. Li, K. A. de Foy, S. N. Cohen, B. A. Ponder, C. Caldas. Aberrant splicing of the TSG101 and FHIT genes occurs frequently in multiple malignancies and in normal tissues and mimics alterations previously described in tumours. *Oncogene*, 15(17), 2119-2126 (1997)
78. Z. Sun, J. Pan, G. Bublely, S. P. Balk. Frequent abnormalities of TSG101 transcripts in human prostate cancer. *Oncogene*, 15(25), 3121-3125 (1997)
79. S. F. Lin, P. M. Lin, T. C. Liu, J. G. Chang, Y. C. Sue, T. P. Chen. Clinical implications of aberrant TSG101 transcripts in acute myeloblastic leukemia. *Leuk Lymphoma*, 36(5-6), 463-466 (2000)
80. T. I. Orban, E. Olah. Emerging roles of BRCA1 alternative splicing. *Mol Pathol*, 56(4), 191-197 (2003)
81. T. I. Orban, E. Olah. Expression profiles of BRCA1 splice variants in asynchronous and in G1/S synchronized tumor cell lines. *Biochem Biophys Res Commun*, 280(1), 32-38 (2001)
82. C. F. Xu, J. A. Chambers, H. Nicolai, M. A. Brown, Y. Hujairat, S. Mohammed, S. Hodgson, D. P. Kelsell, N. K. Spurr, D. T. Bishop, E. Solomon. Mutations and alternative splicing of the BRCA1 gene in UK breast/ovarian cancer families. *Genes Chromosomes Cancer*, 18(2), 102-110 (1997)
83. M. Lu, S. D. Conzen, C. N. Cole, B. A. Arrick. Characterization of functional messenger RNA splice variants of BRCA1 expressed in nonmalignant and tumor-derived breast cells. *Cancer Res*, 56(20), 4578-4581 (1996)
84. M. Mixon, F. Kittrell, D. Medina. Expression of Brcal and splice variant Brcaldelta11 RNA levels in mouse mammary gland during normal development and tumorigenesis. *Oncogene*, 19(46), 5237-5243 (2000)
85. C. A. Wilson, M. N. Payton, G. S. Elliott, F. W. Buaas, E. E. Cajulis, D. Grosshans, L. Ramos, D. M. Reese, D. J. Slamon, F. J. Calzone. Differential subcellular localization, expression and biological toxicity of BRCA1 and the splice variant BRCA1-delta11b. *Oncogene*, 14(1), 1-16 (1997)
86. L. A. Gordon, K. T. Mulligan, H. Maxwell-Jones, M. Adams, R. A. Walker, J. L. Jones. Breast cell invasive potential relates to the myoepithelial phenotype. *Int J Cancer*, 106(1), 8-16 (2003)
87. W. M. ElShamy, D. M. Livingston. Identification of BRCA1-IRIS, a BRCA1 locus product. *Nat Cell Biol*, 6(10), 954-967 (2004)
88. E. Nakuci, S. Mahner, J. Drenzo, W. M. ElShamy. BRCA1-IRIS regulates cyclin D1 expression in breast cancer cells. *Exp Cell Res*, 312(16), 3120-3131 (2006)
89. B. Xu, A. H. O'Donnell, S. T. Kim, M. B. Kastan. Phosphorylation of serine 1387 in Brcal is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation. *Cancer Res*, 62(16), 4588-4591 (2002)
90. R. I. Yarden, S. Pardo-Reoyo, M. Sgagias, K. H. Cowan, L. C. Brody. BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage. *Nat Genet*, 30(3), 285-289 (2002)
91. K. Somasundaram, H. Zhang, Y. X. Zeng, Y. Houvras, Y. Peng, G. S. Wu, J. D. Licht, B. L. Weber, W. S. El-Deiry. Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor p21WAF1/Cip1. *Nature*, 389(6647), 187-190 (1997)
92. M. Santarosa, A. Viel, M. Boiocchi. Splice variant lacking the transactivation domain of the BRCA2 gene and mutations in the splice acceptor site of intron 2. *Genes Chromosomes Cancer*, 26(4), 381-382 (1999)
93. J. P. Zou, Y. Hirose, H. Siddique, V. N. Rao, E. S. Reddy. Structure and expression of variant BRCA2a lacking the transactivation domain. *Oncol Rep*, 6(2), 437-440 (1999)

## Pre-mRNA splicing aberrations and cancer

94. J. Milner, B. Ponder, L. Hughes-Davies, M. Seltmann, T. Kouzarides. Transcriptional activation functions in BRCA2. *Nature*, 386(6627), 772-773 (1997)
95. I. Bieche, R. Lidereau. Increased level of exon 12 alternatively spliced BRCA2 transcripts in tumor breast tissue compared with normal tissue. *Cancer Res*, 59(11), 2546-2550 (1999)
96. V. Staalesen, J. Falck, S. Geisler, J. Bartkova, A. L. Borresen-Dale, J. Lukas, J. R. Lillehaug, J. Bartek, P. E. Lonning. Alternative splicing and mutation status of CHEK2 in stage III breast cancer. *Oncogene*, 23(52), 8535-8544 (2004)
97. S. Agrawal, C. Eng. Differential Expression of Novel Naturally Occurring Splice Variants of PTEN and their Functional Consequences in Cowden syndrome and Sporadic Breast Cancer. *Hum Mol Genet* (2006)
98. K. F. Melo, R. M. Martin, E. M. Costa, F. M. Carvalho, A. A. Jorge, I. J. Arnhold, B. B. Mendonca. An unusual phenotype of Frasier syndrome due to IVS9 +4C>T mutation in the WT1 gene: predominantly male ambiguous genitalia and absence of gonadal dysgenesis. *J Clin Endocrinol Metab*, 87(6), 2500-2505 (2002)
99. D. Wilhelm, C. Englert. The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. *Genes Dev*, 16(14), 1839-1851 (2002)
100. G. B. Silberstein, K. Van Horn, P. Strickland, C. T. Roberts, Jr., C. W. Daniel. Altered expression of the WT1 wilms tumor suppressor gene in human breast cancer. *Proc Natl Acad Sci U S A*, 94(15), 8132-8137 (1997)
101. X. Zhu, A. A. Daffada, C. M. Chan, M. Dowsett. Identification of an exon 3 deletion splice variant androgen receptor mRNA in human breast cancer. *Int J Cancer*, 72(4), 574-580 (1997)
102. G. Lacerra, H. Sierakowska, C. Carestia, S. Fucharoen, J. Summerton, D. Weller, R. Kole. Restoration of hemoglobin A synthesis in erythroid cells from peripheral blood of thalassemic patients. *Proc Natl Acad Sci U S A*, 97(17), 9591-9596 (2000)
103. S. J. Errington, C. J. Mann, S. Fletcher, S. D. Wilton. Target selection for antisense oligonucleotide induced exon skipping in the dystrophin gene. *J Gene Med*, 5(6), 518-527 (2003)
104. S. Fletcher, K. Honeyman, A. M. Fall, P. L. Harding, R. D. Johnsen, S. D. Wilton. Dystrophin expression in the mdx mouse after localised and systemic administration of a morpholino antisense oligonucleotide. *J Gene Med*, 8(2), 207-216 (2006)
105. P. L. Harding, A. M. Fall, K. Honeyman, S. Fletcher, S. D. Wilton. The influence of antisense oligonucleotide length on dystrophin exon skipping. *Mol Ther*, 15(1), 157-166 (2007)
106. L. Cartegni, A. R. Krainer. Correction of disease-associated exon skipping by synthetic exon-specific activators. *Nat Struct Biol*, 10(2), 120-125 (2003)
107. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, 306(5696), 636-640 (2004)
108. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146), 799-816 (2007)
109. R. Tacke, Y. Chen, J. L. Manley. Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. *Proc Natl Acad Sci U S A*, 94(4), 1148-1153 (1997)
110. R. Reiter, A. Wellstein, A. T. Riegel. An isoform of the coactivator AIB1 that increases hormone and growth factor sensitivity is overexpressed in breast cancer. *J Biol Chem*, 276(43), 39736-39741 (2001)
111. D. A. Haber, R. L. Sohn, A. J. Buckler, J. Pelletier, K. M. Call, D. E. Housman. Alternative splicing and genomic structure of the Wilms tumor gene WT1. *Proc Natl Acad Sci U S A*, 88(21), 9618-9622 (1991)
112. S. B. Lee, D. A. Haber. Wilms tumor and the WT1 gene. *Exp Cell Res*, 264(1), 74-99 (2001)
113. A. Ghosh, D. Stewart, G. Matlaszewski. Regulation of human p53 activity and cell localization by alternative splicing. *Mol Cell Biol*, 24(18), 7987-7997 (2004)
114. S. Courtois, G. Verhaegh, S. North, M. G. Luciani, P. Lassus, U. Hibner, M. Oren, P. Hainaut. DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene*, 21(44), 6722-6728 (2002)
115. O. Ishimoto, C. Kawahara, K. Enjo, M. Obinata, T. Nukiwa, S. Ikawa. Possible oncogenic potential of DeltaNp73: a newly identified isoform of human p73. *Cancer Res*, 62(3), 636-641 (2002)
116. I. Casciano, K. Mazzocco, L. Boni, G. Pagnan, B. Banelli, G. Allemanni, M. Ponzoni, G. P. Tonini, M. Romani. Expression of DeltaNp73 is a molecular marker for adverse outcome in neuroblastoma patients. *Cell Death Differ*, 9(3), 246-251 (2002)
117. A. I. Zaika, N. Slade, S. H. Erster, C. Sansome, T. W. Joseph, M. Pearl, E. Chalas, U. M. Moll. DeltaNp73, a dominant-negative inhibitor of wild-type p53 and TAp73, is up-regulated in human tumors. *J Exp Med*, 196(6), 765-780 (2002)
118. S. Tanaka, K. Sugimachi, H. Saeki, J. Kinoshita, T. Ohga, M. Shimada, Y. Maehara. A novel variant of WISP1 lacking a Von Willebrand type C module overexpressed in scirrhous gastric carcinoma. *Oncogene*, 20(39), 5525-5532 (2001)
119. B. He, M. Mirza, G. F. Weber. An osteopontin splice variant induces anchorage independence in human breast cancer cells. *Oncogene*, 25(15), 2192-2202 (2006)
120. F. Bartel, H. Taubert, L. C. Harris. Alternative and aberrant splicing of MDM2 mRNA in human cancer. *Cancer Cell*, 2(1), 9-15 (2002)
121. I. Sigalas, A. H. Calvert, J. J. Anderson, D. E. Neal, J. Lunec. Alternatively spliced mdm2 transcripts with loss of p53 binding domain sequences: transforming ability and frequent detection in human cancer. *Nat Med*, 2(8), 912-917 (1996)
122. J. Lukas, D. Q. Gao, M. Keshmeshian, W. H. Wen, D. Tsao-Wei, S. Rosenberg, M. F. Press. Alternative and aberrant messenger RNA splicing of the mdm2 oncogene in invasive breast cancer. *Cancer Res*, 61(7), 3212-3219 (2001)
123. J. M. Coulson, J. L. Edgson, P. J. Woll, J. P. Quinn. A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. *Cancer Res*, 60(7), 1840-1844 (2000)
124. C. Ghigna, S. Giordano, H. Shen, F. Benvenuto, F. Castiglioni, P. M. Comoglio, M. R. Green, S. Riva, G.



## Pre-mRNA splicing aberrations and cancer

Biamonti. Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene. *Mol Cell*, 20(6), 881-890 (2005)

125. E. Stickeler, V. J. Mobus, D. G. Kieback, P. Kohlberger, I. B. Runnebaum, R. Kreienberg. Intron 9 retention in gene transcripts suggests involvement of CD44 in the tumorigenesis of ovarian cancer. *Anticancer Res*, 17(6D), 4395-4398 (1997)

**Abbreviations:** ESE:Exonic splice enhancer, ESS:Exonic splice silencer, ISE:Intronic splice enhancer, ISS:Intronic splice silencer, ss:Splice site, hnRNP:Heterogeneous nuclear ribonucleoproteins, SR:Serine-arginine rich , NMD:nonsense-mediated decay, IVS:intervening sequence

**Key Words:** Splicing, Regulation Of Splicing, Aberrant Splicing, Mutation, Cancer, Oncogene, Tumour Suppressor Gene, RNA, Review

**Send correspondence to:** Dr Melissa A Brown, School of Molecular and Microbial Sciences, University of Queensland, St. Lucia, Queensland, Australia, 4072, Tel: 617-3365 4628, Fax: 617-3365 4699, E-mail: melissa.brown@uq.edu.au

<http://www.bioscience.org/current/vol13.htm>