Aberrant and Alternative Splicing in Cancer

Julian P. Venables

University of Newcastle-upon-Tyne, Institute of Human Genetics, International Centre for Life, Central Parkway, Newcastle-upon-Tyne, United Kingdom

Abstract

Pre-mRNA splicing is a sophisticated and ubiquitous nuclear process, which is a natural source of cancer-causing errors in gene expression. Intrinsic splice site mutations of tumor suppressor genes often cause exon-skipping events that truncate proteins just like classical nonsense mutations. Also, many studies over the last 20 years have reported cancer-specific alternative splicing in the absence of genomic mutations. Affected proteins include transcription factors, cell signal transducers, and components of the extracellular matrix. Antibodies against alternatively spliced products on cancer cells are currently in clinical trials, and competitive reverse transcription-PCR across regions of alternative splicing is being used as a simple diagnostic test. As well as being associated with cancer, the nature of the alternative gene products is usually consistent with an active role in cancer; therefore, the alternative splicing process itself is a potential target for gene therapy.

Splice Site Mutations Can Cause Aberrant Splicing and Cancer

Defects in mRNA splicing are an important cause of disease (1–3). The most common form of splicing defects are genomic splice site point mutations, and a recent survey found 29 different p53 splice site mutations in >12 different types of cancer (4). Ninety-nine percent of all exons are flanked by the intrinsic dinucleotides GT and AG at the 5’ and 3’ splice sites respectively, and mutation of these sites usually causes exclusion of the adjacent exon (Fig. 1A), although sometimes splice site mutations have even more drastic consequences, e.g., double exon skipping in MLH1 in hereditary nonpolyposis colorectal cancer (5). More than half of all exon deletions lead to truncation of the encoded protein and as such, mutations in tumor suppressor genes at the invariant intrinsic dinucleotide can act much as classical nonsense mutations. For example a GT to AT 5’ splice site mutation in hSNF5 causes deletion of exon 7, a frame shift, and a truncated reading frame, and this causes infant brain tumors when a “second hit” is provided at the wild-type allele by a deletion (6). Similarly a 3’ splice site AG to AT mutation caused constitutive loss of exon 4 of the APC gene in colorectal to liver metastases, where the wild-type allele was deleted (ref. 7; Fig. 1A).

Mutations in the less conserved splice site consensus away from the invariant dinucleotides tend to lead to partial aberrant splicing, often with a relatively mild phenotype. For example, the most common pathogenic mutation of the ATM gene is linked to breast cancer but is incompletely penetrant and is thought to have originated in Palaeolithic times. This is a mutation at the sixth position of an intron that causes a proportion of transcripts to skip an exon and to be frame shifted leading to a truncated protein (8). The polypyrimidine tract signal associated with the 3’ splice site is another hot spot for mutation (Fig. 1A). An intronic mutation 11 nucleotides upstream of a 3’ splice site in MLH1 knocks out an exon, and this causes truncation of the protein and hereditary nonpolyposis colorectal cancer (9). The last position of exons is also highly conserved, and in a relatively mild retinoblastoma case a G to A mutation at that position in RB1 causes an in frame exon to be skipped in >90% of the mature transcripts from the mutated allele (10).

Another way that mutations can cause aberrant splicing is by the inappropriate creation of cryptic splice site signals. An AA to AG mutation creates a cryptic 3’ splice site that adds 11 nucleotides to BRCA1 mRNA, which encodes a truncated protein in a breast cancer family (11). Another breast cancer-causing mutation is an AT to GT change in the estrogen receptor gene, which creates a cryptic 5’ splice site, deep in an intron. This leads to the insertion of a 69-nucleotide cryptic exon into the reading frame (ref. 12; Fig. 1A). In the central nervous system tumor-disposition gene NF2, a CTAGC to CTAAC mutation creates a consensus branch point adenosine (italicized) that creates a cryptic 106 base truncating exon. Because some normal splicing remains, this mutation is associated with a relatively mild phenotype (13).

The final class of cis-genomic mutations that affect splicing are in the relatively uncharacterised “exonic splicing enhancers.” These are degenerate sequences that are often bound by splicing factors known as SR proteins that help recruit the spliceosome to otherwise suboptimal splice sites (14, 15). In NF1, a double point mutation near the middle of a 174 base exon has been found that causes a high level of in-frame deletion of that exon. Consensus binding sites for the SR proteins SC35 and ASF/SF2 are disrupted by the mutations (16). A silent mutation in the middle of exon 14 of APC, a site of natural alternative splicing of an unknown function, has been found to have enhanced exon skipping in a case of familial adenomatous polyposis (17). Also, in various hereditary nonpolyposis colorectal cancer kindreds, different mutations in codon 659 result in skipping of exon 17 of the mismatch repair protein MLH1. This causes an internal deletion of 31 amino acids that abrogates binding to the other mismatch repair protein PMS2 (18).

Alternative Splicing in Cancer

The definition of alternative splicing is the process whereby identical pre-mRNA molecules are spliced in different ways, and this is important in normal development as a means of creating protein diversity in complex organisms (14, 15). Many alternative splicing events have been noted in human development, especially in the brain and the testes (19, 20). Alternative splicing has also been found to be associated with various diseases including growth hormone deficiency, Frasier syndrome, Parkinson’s disease, cystic fibrosis, retinitis pigmentosa, spinal muscular atrophy, and myotonic dystrophy (1, 2). Also, in cancer, there are examples of every kind of alternative splicing, which include the use of alternative individual splice sites, alternative exons, and alternative introns (Fig. 1, B–G). The most intricate type of alternative splicing involves the mutually exclusive use of alternative exons, and a mutually exclusive splice of the
cytoskeletal protein actinin-4 has been discovered recently to be associated with small cell lung cancer (Fig. 1E). The alternative exon is nearly identical to the exon it replaces and changes just three amino acids in the resulting protein, which has a higher affinity for actin and an altered subcellular location as a result; therefore, this could be the cause of the abnormal cytoskeleton found in small cell lung cancer (21). Cancer is a complex integrated process involving signals that come from the extracellular matrix to the nucleus and back again; therefore, for simplicity, the following examples of alternative splicing in cancer are categorized according to their relatively well-characterized subcellular locations, whether it be nuclear, cytoplasmic, transmembrane, or extracellular.

Alternative Splicing of Transcription Factors

The NRSF transcription-silencing factor has an alternative 50-base exon inserted to produce a truncated protein in small cell lung cancer, and this exon is not used in non-small cell lung cancer, so the extra exon is a potential clinical marker for small cell lung cancer (22). Androgen receptor is a transcription factor with an in-frame alternative splice lacking exon 3, which removes 36 amino acids in-frame from the DNA-binding domain. This form was detected in all of 8 breast tumors but not in any of 5 normal samples (23). Sex hormone-binding globulin is a diffusible hormone carrier in blood but also acts as an intracellular cofactor for estrogen in cervical, ovarian, and uterine cancers. A truncated form lacking the estrogen-binding domain encoded by exon 7 was prevalent in 46 tumors and not in 16 normal endometria. Western blotting demonstrated a significant increase in the truncated form during dedifferentiation from G1 to G3 stage tumors, and a concomitant decrease in the full-length form, implying a splicing switch was in operation (24). Another nuclear hormone coactivator, Amplified in Breast Cancer (AIB1), is alternatively spliced to lack exon 3 in breast cancer; 7 of 8 tumors had greater amounts of this form than 6 normal breast tissue samples. The short product uses a downstream initiation codon and is actually more active than the full-length protein at promoting estrogen receptor-mediated transcription, so it may be a significant cause of breast cancer (25).

Global gene expression changes are induced in cancer by hypomethylation of chromatin, and a possible mechanism for this is alternative splicing of the DNA methyltransferase DNMT3b by exclusion of exon 21 to give a frame-shifted and truncated product. Hypomethylation of DNA correlated with the presence of this frame-shifted variant in a large panel of liver tumors, and this may be a general mechanism of chromatin hypomethylation in cancer (26). PASG is homologous to the chromatin-remodelling protein SNF2, and an alternative 5'/H11032 splice site of PASG that removes 75 bases was used in 27 of 57 acute leukemia tumors but 0 of 8 normal PBM samples (27).
Alternative Splicing and Cell Signaling

Soluble cell signaling adaptor proteins are involved in the spread of cancer, and several cellular factors involved in the transduction of oncogenic signals are alternatively spliced in specific cancers. The neurofibromatosis type 1 (NF1) protein is a neural tumor suppressor that functions in part by inactivating ras oncogene signaling. An NF1 variant with an alternative 63-nucleotide insertion in the reading frame is a weaker tumor suppressor and was quantified at between 70% and 90% of the total message in medulloblastomas and primitive neuroectodermal tumors, whereas it is only a minor variant in normal brain tissue (28). Similarly, rac1 is a small GTPase of the ras family involved in cell adhesion, motility, and cell cycle progression. An alternatively spliced form in colorectal cancer contains an extra exon that inserts 19 amino acids into rac1 to give rac1b, which has profoundly altered biochemical characteristics and could not induce lamellipodia like the default rac1 protein (29, 30). Crk oncogene is an SH2 and SH3-domain–containing adapter protein that is spliced into two different forms (Fig. 1C). CrkI uses an upstream 5' splice site that makes it 170 bases shorter than crkII. As a result, crkI is frame-shifted and lacks the COOH-terminal SH3 domain and tyrosine phosphorylation site. In 5 paired samples crkII was the only form detected in normal tissue, but both forms were equally abundant in glioblastomas. Consistent with its exclusive presence in tumors, crkI but not crkII was found to promote cell migration on fibronectin and to promote invasion by activating PI3K/Akt signaling (31). Syk is a soluble tyrosine kinase that suppresses metastasis in vitro. A short form lacking one exon that encodes a novel nuclear localization sequence failed to prevent metastasis and was frequently found in breast tumors but was not found in matched normal tissue (32).

Alternative Splicing of Transmembrane Proteins

Many examples of cancer-related alternative splicing pertain to proteins that are present at the cell surface, and these offer great potential for the selective therapeutic targeting of tumors (see below). The SVH gene (specific Splicing Variant involved in Hepatocarcinogenesis) encodes a transmembrane protein that was discovered for being up-regulated in liver cancer. However, alternative inclusion of two exons results in expression of four splice variants, and actually it is only one variant that is up-regulated in cancer; in 28 of 46 hepatocellular carcinoma samples, levels of SVH-B and none of the other splice variants were elevated in tumor compared with adjacent tissue (33). SVH-B alone caused tumors when xenotransplanted into nude mice, and antisense inhibition of SVH-B caused apoptosis in hepatoma cells. It will be interesting to hear more about the mechanism in tumorigenesis of this novel alternatively spliced gene.

Transmembrane receptor proteins are essential for transmitting growth signals from the extracellular matrix, and several are alternatively spliced in cancer. Specific variants of two G-protein–coupled receptors have been found in pancreatic tumors. The growth-inhibitory secretin receptor without exon 3 has a 36 amino acid deletion, and this was the predominant form in 3 of 3 pancreatic tumors but only a minor product in 3 normal samples. The splice variant acts as a dominant negative because it forms heterodimers that cannot bind secretin (34). The cholecystokinin-B/gastrin receptor regulates gastric cell proliferation, and in pancreatic tumors an extra 69 amino acids are inserted into its intracellular loop due to retention of an intron, which is a relatively rare kind of alternative splicing (Fig. 1D). This receptor is constitutively active and stimulates cell growth. U2AF is a ubiquitous splicing factor that is composed of two subunits of 35 and 65 kDa, respectively. U2AF is an essential part of the splicing reaction, because it is responsible for 3' splice site recognition. Cotransfection experiments showed that U2AF35 but not U2AF65 was limiting for correct intron removal from the gastrin receptor pre-mRNA. Consistent with this, mRNA levels of endogenous U2AF35 in pancreatic tumors were about half the level of surrounding healthy tissue, whereas U2AF65 levels were equal (35).

The fibroblast growth factor receptor FGFR1 exhibits alternative splicing of a single exon encoding an extracellular immunoglobulin disulfide loop. The β form lacks this exon and has a higher affinity for fibroblast growth factors, and unlike the full-length FRFR1α form, FGFR1-β caused cancer in nude mice xenografts (36). The FGFR1 β to α ratio correlates with poor prognosis in breast tumors and with malignancy of astrocytomas (37, 38). Malignant astrocytomas also show up-regulation of the ubiquitous RNA-binding protein PTB. Like U2AF, PTB binds polypyrimidine tracts in introns, but it competes with U2AF, and by this means it inhibits splicing. PTB interacts with a region upstream of the FGFR1 α exon and causes its exclusion when overexpressed in cells. Antisense targeting of either PTB or its binding site reduced the FGFR1 β to α ratio produced from a minigene (39). In tumors, an antibody against PTB stained glial but not neural cells, and therefore it may be useful in classifying neural tumor origins (40). An isoform of the insulin receptor tyrosine kinase (IR-A) that lacks 12 amino acids encoded by exon 11 has a higher affinity for insulin-like growth factor (IGFII) than normal insulin receptor and the responses of an IR-A expressing stable cell line to IGFII were more characteristic of the ligand than the receptor itself i.e., more mitogenic than metabolic. In colon cancer all 10 tumors had ≥65% IR-A variant compared with total, whereas it constituted less than two-thirds in normal tissues (41). IR-A is also up-regulated in thyroid cancer, and increased levels correlated with the state of tumor dedifferentiation. Because IGFII itself is up-regulated in many cancers including thyroid cancer, it may form part of a growth stimulatory autocrine loop with the IR-A splice variant (42).

Several proteins involved in cell adhesion are alternatively spliced in cancer. The tumor suppressor and cell adhesion molecule C-CAM1 has a short frame-shifted form lacking exon 7, which encodes its cytoplasmic domain that is implicated in insulin receptor signaling. A highly significant switch to the short form was shown in 51 non-small cell lung cancer samples compared with their matched adjacent normal tissues (43). MUC1 is a high molecular weight cell adhesives glycoprotein involved in metastasis of which the central region contains a polymorphic number of tandem 20 amino acid repeats. Thyroid cancers showed use of a cryptic 3' splice site upstream of this polymorphic region, which adds nine codons compared with the normal thyroid tissue form (44). KAI1/CD82 is a transmembrane glycoprotein that suppresses metastasis in those cancers in which its expression is elevated. A variant lacking the 28 amino acids encoded by exon 7 fails to localize to the cell membrane and was found specifically up-regulated in gastric cancer tumors associated with short survival time (45). Integrins are ubiquitous heterodimeric transmembrane cell-adhesive glycoproteins that also interact with the cytoskeleton. Integrin β1C has a 116-base exon not found in integrin β1A that changes the cytoplasmic COOH terminus of the protein. The β1C form has been shown to inhibit cell proliferation, β1A to promote it, and β1C appears to be down-regulated in endometrial cancer (46).

The most studied alternatively spliced gene in cancer is CD44, which is a transmembrane protein involved in cell-cell adhesion (reviewed in ref. 47). CD44 has >20 known isoforms due to variable incorporation of 10 alternative exons in its proximal extracellular domain. However “standard” CD44 lacking all of the alternative exons is always predominant. The alternative exons are not independently selected but are generally included in tandem blocks (48); notably, two isoforms with extra exons v4–7 or v8–10 have been associated with cancer (Fig. 1G). Much attention has been paid to
possible transacting factors that bind to exon v5, and its inclusion may depend on a balance between the RNA-binding proteins SAM68 and hnRNP A1 (49). Immunohistochemistry with antibodies against the product of exon v6 shows that isoforms such as v4–7 containing variable exon 6 are frequently up-regulated in squamous cell carcinomas (50, 51). A highly promising therapeutic tool is antibody targeting of tumors, and radiolabeled anti-exon v6 antibodies are in clinical trials for treatment of head and neck cancer (51, 52).

Because the alternative splicing of CD44 is so complex, fingerprint analysis of alternative splicing was originally performed by subjecting PCR products from across the variable region to Southern blotting with each of the variable exons as probes. Comparisons using a probe to the nonvariable region showed that variant forms were relatively up-regulated in colorectal cancer and also that metastatic cancers were enriched for the larger variable forms containing more variable exons (53, 54). Also, these larger forms correlated with down-regulation of most SR protein mRNAs apart from the smaller SR proteins of 20 and 30 kDa. Therefore, these small SR proteins may have a role in CD44 splicing and metastasis, and they were also found to be up-regulated in a mouse mammary tumorigenesis model (55). 9G8 is one of the 30-kDa SR proteins, and a 9G8-responsive exonic splicing enhancer sequence was found in the v9 exon (56).

Clinically, blood can be assayed for variable CD44 exon inclusion. Using an anti-v9 peptide antibody on blood from a group of 71 colorectal cancer patients, including 14 with liver metastases, showed the metastatic subset had double the blood levels of variants containing v9 (57). CD44v8–10 and v9–10 were found in non-small cell lung cancer tumors but not in adjacent tissue by reverse transcription-PCR (58). A sensitive PCR strategy was used to only amplify the v8-v10 and v10 isoforms and not the standard form from exfoliated cells in urine; both exon v8 and v10 start with the same three nucleotides, so a primer from the end of the common region ending in the same three nucleotides revealed a v8–10 to v10 ratio above 1 in malignant bladder cancers and below 0.7 otherwise (59). These results have been followed up in a larger study and v8-v10 was detected in most urothelial cancer cases and correlated with disease progression, invasive disease and low survival rate (60). The v10 exon is important for cell adhesion in cancer, because it contains a small motif that binds to chondroitin sulfate moieties that are covalently attached to standard CD44 (61).

**Alternative Splicing of Secreted Extracellular Proteins**

Extracellular proteins are essential for metastatic growth, and the urokinase-type plasminogen activator is a secreted anchored glycoprotein with a role in invasion and metastasis, because it is involved in the degradation of extracellular matrix proteins. In 39 breast cancer samples, those with above the median ratio of double-exon–deleted to wild-type urokinase-type plasminogen activator, as determined by quantitative PCR, had a significantly poorer prognosis than those below the median (62). A splice variant of the Wnt-induced secreted protein (WISP1), which lacks a protein interaction region encoded by exon 3, was found in 18 of 21 scirrhous gastric carcinoma supernatants but not in matched normal stroma. Overexpression of this variant causes invasion of cells through collagen in vitro, whereas the full-length protein could not (63). Vascular endothelial growth factor is a secreted protein that is important for angiogenesis. In non-small cell lung cancer there is a shift from the full-length isoform toward two shorter diffusible forms that lack exon 6 encoding the 41 amino acid heparin-binding domain (64).

Tenascin-C is a secreted oligomeric glycoprotein of the extracellular matrix. Normal tissues express a 6kb splice variant of tenasin-C encoding 8 fibronectin type III (FN3) domains. A larger 8-kb variant has 8 extra exons (Fig. 1F) that encode 7 additional FN3 domains, and only this larger variant can facilitate cell migration by inducing loss of focal adhesion, and it may also protect cancer cells from antitumor immune responses (65). Large forms were detected by Northern blotting in invasive breast cancer and quantified at 60% to 90% of the total, whereas in noninvasive disease long isoforms constituted <25% of the total tenasin-C (66). At the protein level the long form was not as overexpressed as expected from the RNA studies in breast cancer, but an antibody against the extra region of tenasin-C did specifically detect glioblastoma tissue in the brain so it may have use as a scintographic marker; however, this antibody has not been used for radiotherapy, because it also targets some extracranial normal tissues (67).

The most promising therapeutic target for antibodies is fibronectin itself, which is also a secreted extracellular protein, and a key determinant of tumor cell control of proliferation, migration, invasion, and metastatic behavior. Alternative splicing of fibronectin involves three extra domains, one of which, the extra domain-B region encodes an inserted 91 amino acids that are 100% conserved in mammals, and this region has been consistently found in cancer. An extra domain-B–specific antibody was used to diagnose high-grade astrocytoma more accurately than any previous methods (68), and trials have now begun to use the antibody to target radioactive iodine to colorectal and lung tumors (69, 70). The fibronectin extra domain-B region was found to have a purine-rich splicing-enhancer that responds to the 40-kDa SR protein (71), and overexpression of SRp40 was shown to occur during mouse mammary tumorigenesis; however, to date no systematic study of the relative expression of different transacting splicing factors in different types of human cancer has been carried out to understand the causes of alternative splicing in cancer (55).

**Does Alternative Splicing Cause Cancer?**

It is difficult to prove that any process is a cause of cancer as it progresses in several stages. First, a progenitor cell is transformed by becoming insensitive and resistant to its surroundings, and then it initiates aggressive interactions leading to angiogenesis, invasion, and metastasis. In some cases cancer-associated splice forms are accompanied by other molecular changes. For example, alternative splicing of cSrc was found to be a favorable indicator in childhood brain tumors, but it correlated with the absence of myc amplification, which is already known to be associated with aggressive cancer (72). Similarly, human T-cell leukemia virus is known to cause adult T-cell leukemia, and a splicing switch from fynT to fynB (Fig. 1E) could be induced by overexpression of the virally encoded rex protein, so it is not clear whether alternative splicing of fyn is a by-product or an etiological factor in adult T-cell leukemia (73). However, in many simple examples alternative splicing of genes involved in these processes correlates with specific kinds of human cancer, and the function of the alternative form is consistent with a possible role in cancer (Table 1).

To be a cause of cancer an alternative splice must presumably be expressed at significant levels compared with properly spliced product. The MDM2 oncoprotein inhibits apoptosis by targeting the tumor suppressor p53 for degradation by the proteasome; however, MDM2 splice variants without an intact NH2-terminal p53-binding domain could still cause transformed cell foci on confluent cell cultures implying that MDM2 has both p53-dependent and p53-independent mechanisms of action (74). The main alternative splice form is MDM2b, which cannot bind p53 because it is missing 8 exons that include the p53-binding domain (Fig. 1F). MDM2b PCR products were exclusively detected in osteosarcomas and not normal bone and...
were enhanced in advanced astrocytomas, ovarian cancers, and bladder cancers (74–76). One possible way that MDM2b could interfere with full-length MDM2 function is by sequestering it in the cytoplasm (77); however, MDM2b must be expressed at a low level, because it has only been detected by reverse transcription-PCR, which will naturally favor this product, because it is 919 bp shorter than full length.

Some alternative splices may be associated with the stresses incurred by diseases, but they may not be specifically associated with cancer. For example CDC44v6 is found in some nondoncancerous thyroid diseases as well as in thyroid cancer (78). The TSG101 tumor suppressor is also alternatively spliced in many cancers, and the delta154–1054 splice variant (Fig. 1) is up-regulated in advanced breast and cervical cancer (79–82). Interestingly, stresses such as hypoxia and cobalt chloride could also induce the delta154–1054 isoform (81).

It is hard to explain how very specific alternative splices are formed and that some spliced variants are selectively mediated. For example, detection is hindered by the fact that alternative splicing is rarely an all or nothing event, and many interesting alternative splices are missed because they are “normal” in other tissues (21); however, one can still study tumor-specific splicing patterns from individual cancers, which may not always be equivalent to primary cancers; for example, alternative splicing of caspase 8 (and resistance to apoptosis) is lost on the establishment of adult T-cell leukemia cell lines (90). Also, tumor-specific splicing patterns from individual cancers may be missed because they are “normal” in other tissues (21); however, one can still study tumor-specific splicing patterns from individual cancers, which may not always be equivalent to primary cancers; for example, in colorectal cancer, progesterone receptor is aberrantly spliced by the apparently random removal of various exons, as well as by utilization of an alternative downstream 3’ splice site (83). Estrogen receptor α also has various missing exons and more double-exon–lacking variants in breast cancer. Many of the aberrant forms are predicted to be estrogen-independent and constitutively active; therefore, alternative splicing may confer a selective advantage to cancer cells without being the primary cause of malignant transformation (84). It may therefore be an oversimplification to ask whether alternative splicing is a cause or an effect of cancer. An additional layer of complexity is that multiple types of cancer can be studied in parallel. One way of studying the association of alternative splicing with cancer at a global level is through the wealth of expressed sequence tag information in the public databases. Several bioinformatics laboratories have studied differential alternative splicing between cancer and normal expressed sequence tag libraries (86–88), and there appears to be a higher proportion of disrupted spliced variants in tumor suppressor genes than in noncancer-related genes (87). The current expressed sequence tag libraries are made in different ways from one another and contain many redundant sequences mostly from abundant genes and 3’ ends (89). Many “cancer” libraries are from cell lines, which may not always be equivalent to primary cancers; for example, alternative splicing of caspase 8 (and resistance to apoptosis) is lost on the establishment of adult T-cell leukemia cell lines (90). Also, tumor-specific splicing patterns from individual cancers may be missed because they are “normal” in other tissues (21); however, one bioinformatics study was successfully validated by competitive reverse transcription-PCR with a heterogeneous group of tumor samples implying that theoretical and experimental approaches are complementary (88).

A recent approach to study alternative splicing is the Exon Junction Microarray, which currently has 125,000 36mer oligonucleotides, corresponding to the exon to exon junctions from 10,000 genes, printed on it. This has been used to discover many alternative splicing events in low abundance genes that are hard to find by expressed sequence tag analysis. A major advantage of this experimental technique is that multiple types of cancer can be studied in parallel. However, detection is hindered by the fact that alternative splicing is rarely an all or nothing event, and many interesting alternative splices are rapidly degraded by nonsense-mediated decay (91).

The existence of >30,000 genes with multiple exons that could be alternatively spliced in any of >100 types of cancer means that the many examples of cancer-related alternative splicing thus far obtained (experimentally and bioinformatically) are likely to be the tip of the iceberg.

Global Studies of Splicing Changes in Cancer

One way of studying the association of alternative splicing with cancer at a global level is through the wealth of expressed sequence tag information in the public databases. Several bioinformatics laboratories have studied differential alternative splicing between cancer and normal expressed sequence tag libraries (86–88), and there appears to be a higher proportion of disrupted spliced variants in tumor suppressor genes than in noncancer-related genes (87). The current expressed sequence tag libraries are made in different ways from one another and contain many redundant sequences mostly from abundant genes and 3’ ends (89). Many “cancer” libraries are from cell lines, which may not always be equivalent to primary cancers; for example, alternative splicing of caspase 8 (and resistance to apoptosis) is lost on the establishment of adult T-cell leukemia cell lines (90). Also, tumor-specific splicing patterns from individual cancers may be missed because they are “normal” in other tissues (21); however, one bioinformatics study was successfully validated by competitive reverse transcription-PCR with a heterogeneous group of tumor samples implying that theoretical and experimental approaches are complementary (88).

A recent approach to study alternative splicing is the Exon Junction Microarray, which currently has 125,000 36mer oligonucleotides, corresponding to the exon to exon junctions from 10,000 genes, printed on it. This has been used to discover many alternative splicing events in low abundance genes that are hard to find by expressed sequence tag analysis. A major advantage of this experimental technique is that multiple types of cancer can be studied in parallel. However, detection is hindered by the fact that alternative splicing is rarely an all or nothing event, and many interesting alternative splices are rapidly degraded by nonsense-mediated decay (91).

The existence of >30,000 genes with multiple exons that could be alternatively spliced in any of >100 types of cancer means that the many examples of cancer-related alternative splicing thus far obtained (experimentally and bioinformatically) are likely to be the tip of the iceberg.

### Table 1 Selected examples of cancer-specific alternative splicing categorized by affected tissue

<table>
<thead>
<tr>
<th>Cancer tissue</th>
<th>Gene</th>
<th>Function</th>
<th>Transacting factor</th>
<th>Selected references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leukemia</td>
<td>Fyn</td>
<td>Tyrosine kinase</td>
<td>Rex</td>
<td>(73)</td>
</tr>
<tr>
<td>Leukemia</td>
<td>Caspase 8</td>
<td>Apoptosis</td>
<td>(90)</td>
<td></td>
</tr>
<tr>
<td>Leukemia</td>
<td>PASG</td>
<td>Chromatin modelling</td>
<td>(27)</td>
<td></td>
</tr>
<tr>
<td>Thyroid</td>
<td>MUC1</td>
<td>Adhesion, metastasis</td>
<td>(44)</td>
<td></td>
</tr>
<tr>
<td>Thyroid, colon</td>
<td>Insulin receptor</td>
<td>Tyrosine kinase</td>
<td>(41, 42)</td>
<td></td>
</tr>
<tr>
<td>Colorectal</td>
<td>Rac1</td>
<td>Signalling GTPase</td>
<td>(29, 30)</td>
<td></td>
</tr>
<tr>
<td>Gastric</td>
<td>KAI1/CD82</td>
<td>Metastasis</td>
<td>(45)</td>
<td></td>
</tr>
<tr>
<td>Gastric</td>
<td>WEER1</td>
<td>Invasion</td>
<td>(63)</td>
<td></td>
</tr>
<tr>
<td>Pancreas</td>
<td>Secretin receptor</td>
<td>Growth inhibitor</td>
<td>(34)</td>
<td></td>
</tr>
<tr>
<td>Pancreas</td>
<td>Gastrin receptor</td>
<td>Proliferation</td>
<td>U2AF (35)</td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>DNMT3b4</td>
<td>Chromatin modelling</td>
<td>(26)</td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>SVH</td>
<td>Novel</td>
<td>(33)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>NR5F</td>
<td>Transcription factor</td>
<td>(22)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>C-CAM</td>
<td>Adhesion</td>
<td>(43)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>VEGF</td>
<td>Angiogenesis</td>
<td>(64)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>Actnin-4</td>
<td>Adhesion, metastasis</td>
<td>(21)</td>
<td></td>
</tr>
<tr>
<td>Endometrium</td>
<td>SHBG</td>
<td>Hormone signalling</td>
<td>(24)</td>
<td></td>
</tr>
<tr>
<td>Endometrium</td>
<td>Integrin β1C</td>
<td>Adhesion</td>
<td>(46)</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>AIB1</td>
<td>Hormone signalling</td>
<td>(25)</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>Androgen receptor</td>
<td>Transcription Factor</td>
<td>(23)</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>Estrogen receptor</td>
<td>Transcription Factor</td>
<td>(84)</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>Syk</td>
<td>Metastasis</td>
<td>(32)</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>uPAR</td>
<td>Adhesion, proteolysis</td>
<td>(62)</td>
<td></td>
</tr>
<tr>
<td>Breast, brain</td>
<td>FGFR1</td>
<td>Growth signalling</td>
<td>FTB (39)</td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>Cdk</td>
<td>Migration, invasion</td>
<td>(31)</td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>NF1</td>
<td>Signalling GTPase</td>
<td>(28)</td>
<td></td>
</tr>
<tr>
<td>Many</td>
<td>TSG101</td>
<td>Proteolysis</td>
<td>(79, 81, 82)</td>
<td></td>
</tr>
<tr>
<td>Many</td>
<td>MDM2</td>
<td>Proteolysis</td>
<td>(74–76)</td>
<td></td>
</tr>
<tr>
<td>Many</td>
<td>CD44</td>
<td>Proliferation, angiogenesis</td>
<td>9G8, SAM68 (47, 49, 51)</td>
<td></td>
</tr>
<tr>
<td>Many</td>
<td>Tenascin-C</td>
<td>Adhesion inhibitor</td>
<td>(66)</td>
<td></td>
</tr>
<tr>
<td>Many</td>
<td>Fibronectin</td>
<td>Angiogenesis</td>
<td>SRp40 (69)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Cancer tissue, gene name, function, and transacting factors known to cause the alternative splice and selected references are in the columns shown.
iceberg, and what is needed to complete the acquisition of alternatively spliced sequences for computational statistical analysis is a systematic method of enriching alternatively spliced isoforms from cDNA libraries. Enrichment of alternatively spliced isoforms relies on their inability to self-hybridize in the alternatively spliced region, and development of novel single-stranded DNA binding technology looks set to improve this method (92).

Alternative Splicing and Cancer Treatment

As well as the direct targeting of alternatively spliced protein isoforms, such as fibronectin extra-domain-B and CD44v6, dealt with in the previous sections (51, 69), there are several methods of therapeutic intervention in cancer that either alter or exploit the alternative splicing process itself. Methods differ according to whether drugs target transcripts directly or work by affecting trans-acting splicing factors. There are several direct antisense-based methods being developed to remedy aberrations of splicing in genetic disease including stable antisense, RNA interference, and hybrid protein nucleic acids (reviewed in ref. 2).

There are many examples of alternative splicing changes involved in apoptosis (93), so altering alternative splicing can be used to selectively kill cancer cells. BclX is a member of the Bcl family that has an important role in the breakdown of mitochondria during apoptosis, and it is alternatively spliced between a long antiapoptotic form (BclXl) and a short apoptosis-promoting form (BclXs), which is made by use of an upstream 5′ splice site (Fig. 1C). The downstream 5′ splice site of BclX can be blocked directly (but not degraded) by stable antisense oligonucleotides that divert splicing toward the proapoptotic upstream 5′ splice site. This is a particularly powerful approach, because the more antiapoptotic isoform a cancer cell has, the more cytotoxic proapoptotic form can be made by switching splice site usage (94).

Several agents are thought to affect BclX splicing at the level of transacting splicing factors. A 53 amino acid segment from the interleukin 1α protein was found to divert splicing toward the proapoptotic form. This peptide may be a useful anticancer drug, because it caused apoptosis in many malignant but not normal cell lines, and consistent with a possible direct function in alternative splicing, it retrieved several known splicing factors in a yeast two-hybrid screen (95). The lipid ceramide can also cause the proapoptotic BclX splicing switch. Ceramide inhibits PP1, which dephosphorylates the splicing factor ASF/SF2, and the splicing switch depends on an ASF/SF2 binding site near the BclXs splice site (96). The anticancer drug NB-506, on the other hand, is thought to function in part by inhibiting phosphorylation of ASF/SF2 (97).

An additional method of splicing-related gene therapy relies on exploiting the fact that alternative splices can be cancer-specific, whereas there are only a limited number of promoters that can drive gene expression specifically in cancer cells. In this study cells were rendered sensitive to the antitumor prodrug etoposide phosphate by a minigene expression vector containing alkaline phosphatase downstream of the CD44 v9–10 intron. Alkaline phosphate is only expressed in frame if the intron is spliced out, and this only occurred in cells that endogenously express the cancer-specific CD44v9–8–10 isoform (Fig. 1G; ref. 98). Therefore, this method offers exciting opportunities to specifically target tumors with cytotoxic agents.

In conclusion, there is now ample evidence that just as alternative splicing is important for differentiation, so aberrations of alternative splicing are important for cancer. Various strategies are currently being used to exploit alternative splicing for the diagnosis and treatment of cancer. Potential therapeutic targets that need additional exploration include the trans-acting factors that cause alternative splicing in cancer cells. Equally, we are at an early stage in characterizing the full repertoire of cancer-associated alternatively spliced isoforms. However, this promises to provide a qualitative fingerprint of gene expression in cancer. Whereas alternative splicing in cancer has long been recognized, its significance in the diagnosis and treatment of cancer is just now coming to the fore (99, 100), and many exciting developments in this field seem likely in the foreseeable future.

References


SPEICING IN CANCER


Aberrant and Alternative Splicing in Cancer

Julian P. Venables


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/64/21/7647

Cited articles
This article cites 99 articles, 40 of which you can access for free at:
http://cancerres.aacrjournals.org/content/64/21/7647.full#ref-list-1

Citing articles
This article has been cited by 64 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/64/21/7647.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.