Alterations of pre-mRNA splicing in cancer

ARTICLE *in* GENES CHROMOSOMES AND CANCER · MAY 2005

Impact Factor: 4.04 · DOI: 10.1002/gcc.20156 · Source: PubMed

4 AUTHORS, INCLUDING:

Zane Kalnina
Latvian Biomedical Research and Study Ce…
29 PUBLICATIONS **430** CITATIONS
SEE PROFILE

Pawel Zayakin
Latvian Biomedical Research and Study Ce…
18 PUBLICATIONS **239** CITATIONS
SEE PROFILE

Aija Linē
Latvian Biomedical Research and Study Ce…
17 PUBLICATIONS **413** CITATIONS
SEE PROFILE

GENES, CHROMOSOMES & CANCER 42:342–357 (2005)

REVIEW ARTICLE

Alterations of Pre-mRNA Splicing in Cancer

Zane Kalniņa, Pawel Zayakin, Karīna Siliņa, and Aija Linē*

Biomedical Research and Study Centre, University of Latvia, Riga, Latvia

Recent genomewide analyses of alternative splicing (AS) indicate that up to 70% of human genes may have alternative splice forms, suggesting that AS together with various posttranslational modifications plays a major role in the production of proteome complexity. Splice-site selection under normal physiological conditions is regulated in the developmental stage in a tissue type-specific manner by changing the concentrations and the activity of splicing regulatory proteins. Whereas spliceosomal errors resulting in the production of aberrant transcripts rarely occur in normal cells, they seem to be an intrinsic property of cancer cells. Changes in splice-site selection have been observed in various types of cancer and may affect genes implicated in tumor progression (for example, *CD44, MDM2,* and *FHIT*) and in susceptibility to cancer (for example, *BRCA1* and *APC*). Splicing defects can arise from inherited or somatic mutations in *cis*-acting regulatory elements (splice donor, acceptor and branch sites, and exonic and intronic splicing enhancers and silencers) or variations in the composition, concentration, localization, and activity of regulatory proteins. This may lead to altered efficiency of splice-site recognition, resulting in overexpression or down-regulation of certain splice variants, a switch in splice-site usage, or failure to recognize splice sites correctly, resulting in cancer-specific splice forms. At least in some cases, changes in splicing have been shown to play a functionally significant role in tumorigenesis, either by inactivating tumor suppressors or by gain of function of proteins promoting tumor development. Moreover, cancer-specific splicing events may generate novel epitopes that can be recognized by the host's immune system as cancer specific and may serve as targets for immunotherapy. Thus, the identification of cancer-specific splice forms provides a novel source for the discovery of diagnostic or prognostic biomarkers and tumor antigens suitable as targets for therapeutic intervention.

CONSTITUTIVE AND ALTERNATIVE SPLICING

Basal Splicing Machinery

The coding regions (exons) of most human genes are interrupted by noncoding intervening sequences (introns) that are removed from pre-mRNA molecules to produce mature mRNAs through the process of RNA splicing. RNA splicing is carried out cotranscriptionally by a spliceosome—a multicomponent complex consisting of five small nuclear ribonucleoproteins (snRNPs—U1, U2, U4, U5, and U6) and more than 100 proteins (Faustino et al., 2003). Via multiple RNA–RNA, RNA–protein, and protein–protein interactions, the spliceosome recognizes exon–intron boundaries and catalyzes two sequential trans-esterification reactions that remove introns and ligate exons (Patel et al., 2003). Intron removal must be performed with the precision of up to one nucleotide. However, it is not fully understood how it is accomplished because exon–intron borders are defined by short and weakly conserved classical splice-site sequences (Fig. 1). The 5′ and the 3′ termini of an intron are defined by a GU and an AG dinucleotide, respectively, in 95%–99% of cases (Burset et al., 2000; Black, 2003). In the 5′ splice site, the GU dinucleotide is followed by a less conserved consensus sequence, whereas the 3′ splice site consists of a branch point most often containing an adenine residue, followed by the pyrimidine-rich region and the conserved AG dinucleotide (Black, 2003). In addition to these sequences, exonic and intronic *cis* elements, known as splicing enhancers and silencers, are involved in both constitutive and alternative splicing, in which they aid in the correct identification of exon–intron borders and prevent pseudoexons from being included in the mRNA. They appear to be particularly important in the regulation of splice-site usage during alternative splicing (Cartegni et al., 2002).

Alternative Splicing

Alternative splicing (AS) of a pre-mRNA is a fundamental mechanism of differential gene expression that allows the production of structurally and functionally distinct proteins from a single coding

Supported by: Latvian Council of Science and the Royal Society.

*Correspondence to: Aija Linē, Biomedical Research and Study Centre, University of Latvia, Ratsupites St 1, LV-1067, Riga, Latvia. E-mail: aija@biomed.lu.lv

Received 29 July 2004; Accepted 4 November 2004

DOI 10.1002/gcc.20156

Published online 11 January 2005 in Wiley InterScience (www.interscience.wiley.com).

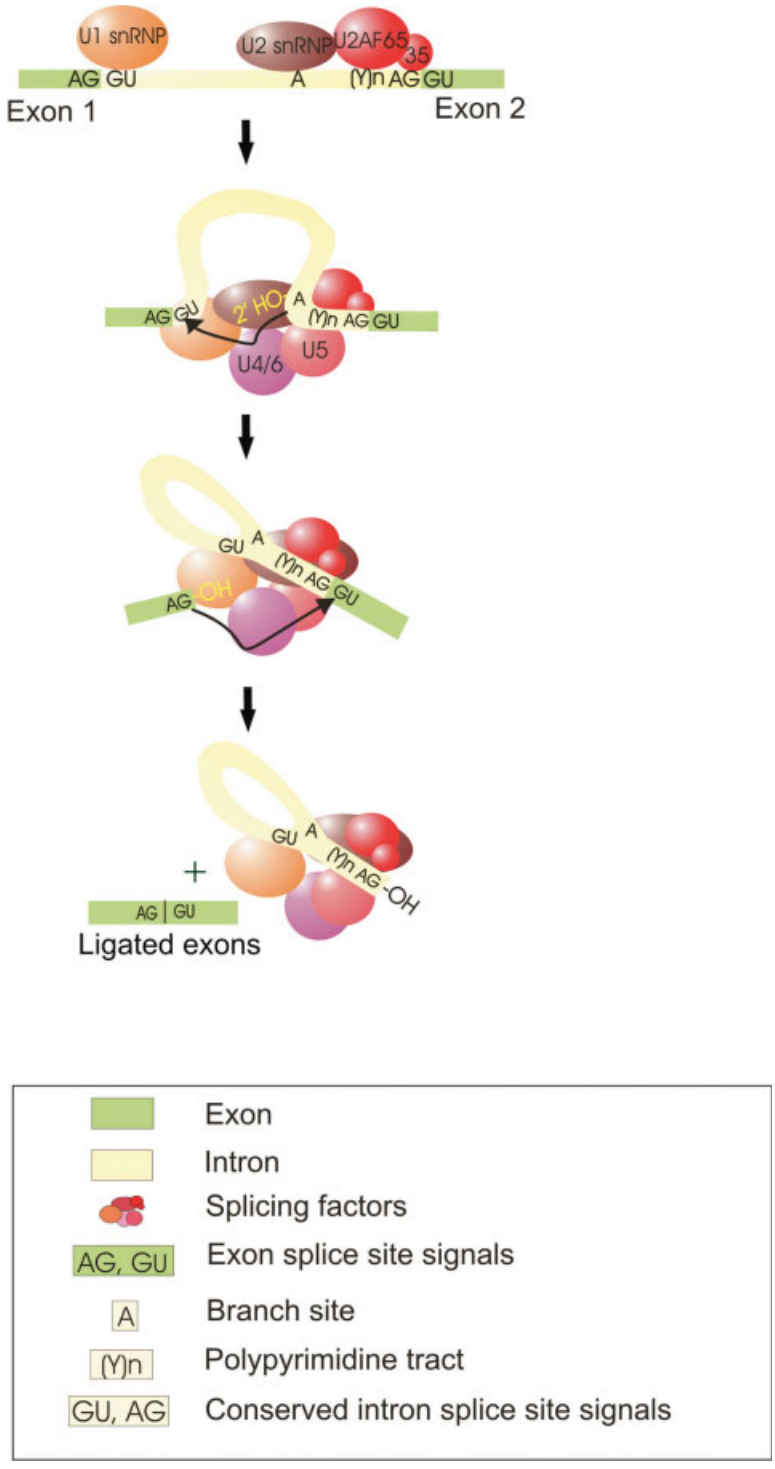

Figure 1. Molecular mechanism of splicing. Assembly of spliceosome begins with the base-pairing interactions of U1 and U2 snRNPs with the 5′ splice site and branch site, respectively. Then the U4–U6 complex joins the prespliceosome and, after base-pairing interactions with the pre-mRNA, displaces the U1 and U4 snRNPs. In the first transesterification reaction, the 2′ OH group of the adenosine residue at the branch site attacks the 5′ splice site, generating a free 3′ OH group on the first exon and a branched lariat. Next, the 5′ nucleotide of the downstream exon is attacked by the free 3′ OH group, cleaving the RNA molecule at the 3′ splice site. The exon sequences are ligated to each other, and the intron is released as a lariat (Patel et al., 2003).

sequence. The finding that the human genome contains only about 30,000–40,000 genes—only about twice as many as in the fly (Lander et al., 2001)—came as a surprise to many. Nevertheless, the proteome might number more than several hundred thousand protein species. Numerous genomewide analyses of AS have indicated that 35%–74% of human genes may have alternative splice forms (Modrek et al., 2001, 2002; Johnson et al., 2003), suggesting that AS together with various posttranslational modifications plays a major role in the production of proteome complexity.

At least five distinct alternative splicing patterns have been observed (Fig. 2): (1) the cassette exon

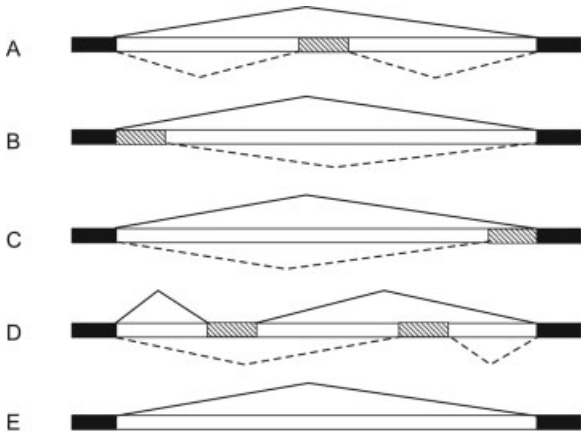

Figure 2. Patterns of alternative splicing: (A) exon skipping/inclusion; (B) alternative 5′ splice site; (C) alternative 3′ splice site; (D) mutually exclusive exons; (E) intron retention (Cartegni et al., 2002; Black, 2003).

that can either be skipped or be included in the mRNA; (2) the alternative 5′ splice site; (3) the alternative 3′ splice site; (4) the complex splicing pattern, when only one of the two mutually exclusive cassette exons is included in the mRNA; and (5) intron retention. Different types of alternative splicing can be combined to generate multiple mRNA isoforms from a single gene (Cartegni et al., 2002; Black, 2003). Besides that, different 5′ and 3′ ends of mRNA also can be generated by the use of alternative promoters in different exons and alternative polyadenylation signals.

Splicing Enhancers and Silencers

Exonic and intronic splicing enhancers (ESEs and ISEs) and silencers (ESSs and ISSs) are *cis*-regulatory elements that in most cases serve as binding sites for splicing factors, which stimulate or repress, respectively, splice-site usage or create an RNA secondary structure that affects splice-site recognition (Maniatis et al., 2002; Black, 2003). Frequently, enhancers and silencers overlap to form a composite regulatory element. ESEs are recognized by SR proteins, so far the best-studied splicing factors, which contain one or two RNA-binding domains and an arginine-serine (RS)–rich domain. The latter is thought to stimulate spliceosome assembly by recruiting U1 snRNP and U2AF via protein–protein interactions (Cartegni et al., 2002). However, recently it was also shown to interact directly with the pre-mRNA branch point (Shen et al., 2004). Moreover, SR proteins also can modulate splice-site recognition by competing with negative regulatory factors, but some may act as splicing repressors (Zhu et al., 2001; Cartegni et al., 2002; Maniatis et al., 2002).

ESEs have been identified through the analysis of disease-associated sequence variations (Pagani et al., 2004), by the in vivo or in vitro SELEX (systematic evolution of ligands by exponential enrichment) technique, which is based on the selection of sequences that promote splicing from a pool of random sequences (Cartegni et al., 2002), and by computational approaches based on statistical analysis of exon–intron and splice-site composition (Fairbrother et al., 2002). ESE consensus sequences have been derived from known SR protein binding sites, and the score matrices for individual SR proteins were calculated according to nucleotide frequencies at each position (Liu et al., 1998, 2000). SNPs that affect SR protein-binding scores in some cases have been shown to change splicing efficiency (Liu et al., 2001).

ESSs and ISSs are recognized by hnRNPs, a large group of diverse pre-mRNA binding proteins. Their mechanism of action is less well understood than that of enhancers. In some cases, enhancers and silencers partially overlap, and then the splicing efficiency is determined by the antagonistic action of hnRNPs and SR proteins. hnRNPs bound to ISSs flanking an alternative exon also may cause the exon to loop out, resulting in skipping of the exon. Alternatively, inhibitory factors bound to ESS may polymerize along the exon and displace the ESE-bound SR proteins (Zhu et al., 2001; Cartegni et al., 2002; Maniatis et al., 2002).

Intronic regulatory sequences often are within 50 bp of the splice sites. However, some regulatory elements are found thousands of bases from the splice sites. Some ISEs also are recognized by SR proteins, whereas others bind SR nonrelated proteins (Black, 2003; Pagani et al., 2004).

Regulation of Splice Site Selection

The selection of alternative splice sites can be regulated in different manners: tissue specificity, developmental stage, physiological processes, sex determination, and in response to various stress factors. A number of signals, including stimulation of receptors by growth factors, cytokines, or hormones; depolarization; rising intracellular Ca^{2+} levels; and cellular stresses like heat shock and change in pH, have been shown to induce changes in the selection of splice sites (Scotet et al., 1998; Stamm, 2002; Faustino et al., 2003). Splice-site selection in stress responses or in varying physiological environments can be regulated through changing the activity or composition of general

splicing factors by regulating transcription, by phosphorylation, or by changing their intracellular localization (Stamm, 2002). For example, insulin stimulates the inclusion of the B2 exon of protein kinase C (PKC), resulting in production of the PKCB2 isozyme. The binding of insulin to its receptor activates the phosphatidylinositol 3-kinase signaling pathway, leading to phosphorylation of the SR protein SRp40, which then promotes incorporation of the B2 exon (Patel et al., 2004). CD45, a transmembrane protein tyrosine phosphatase, is alternatively spliced in response to T-cell activation. The switch in CD45 splicing has been shown to be mediated by changing the expression level of several SR proteins, including SC35 (ten Dam et al., 2000; Wang et al., 2001a). In addition to general splicing factors present in multiple tissues, several cell-type specific splicing factors that regulate the splicing of specific subsets of genes have been identified. For example, the *Drosophila* halfpint protein (Hfp) regulates a number of genes required for oogenesis (Van Buskirk et al., 2002). So far, tissue-specific splicing factors in mammals predominantly have been found in the nervous system. One of the best-characterized examples is NOVA1, a neuron-specific RNA-binding protein that has been shown to stimulate the incorporation of exon E9 of *GABA(A)Rgamma2* by binding to an intronic splicing enhancer (Dredge et al., 2001, 2003).

METHODS FOR IDENTIFICATION AND ANALYSIS OF ALTERNATIVE SPLICE VARIANTS

So far, the most reliable technique for the identification and confirmation of alternative splice variants has been the sequencing of full-length cDNA. Various RT-PCR-, RACE (rapid amplification of cDNA ends)–, and Northern blot–based techniques are also widely used for the identification of alternative splice events. However, the power of these approaches for the characterization of alternative splicing patterns across all tissues, developmental stages, or disease conditions is limited. Recently, several approaches for the genome-wide analysis of AS have emerged, including various bioinformatic approaches based on the alignment of mRNAs and expressed sequence tags (ESTs) to the corresponding genomic sequence (Kan et al., 2001; Modrek et al., 2001, 2002; Xie et al., 2002; Xu et al., 2002), exon junction oligonucleotide microarrays (Johnson et al., 2003), AS profiling using a bead-based fiber-optic microarray platform (Yeakley et al., 2002), and polymerase colony (polony) technology (Mitra et al., 2003; Zhu et al., 2003).

Bioinformatic Approaches

More than 5.49 million ESTs derived from more than 6,000 human cancerous- and normal-tissue cDNA libraries have been deposited in the dbEST (http://www.ncbi.nlm.nih.gov/dbEST/) as of this writing (dbEST release 050704, May 7, 2004). These sequences provide a broad sample of mRNA diversity. However, it should be taken into account that genes with lower expression have a coverage bias and that all human tissues and developmental stages are not equally represented, with ESTs from brain and cell line libraries the most abundant. It is also a question whether alternative splice isoforms found in cell lines are functionally representative of those in the human organism. Nevertheless, EST data sets make a valuable source of information and, together with the available sequence of the human genome, have enabled the development of various bioinformatic approaches for the identification of alternative splice variants. Currently, bioinformatics is the only approach capable of identifying all types of novel AS forms on a genomewide scale. Most studies rely on the alignment of EST and mRNA sequences to each other (Wang et al., 2003) or to the genomic sequence (Kan et al., 2001; Modrek et al., 2001; Xie et al., 2002; Thanaraj et al., 2004). In the first step, clusters of ESTs corresponding to the same genomic locus are obtained either by a BLAST search of genomic (Kan et al., 2001) or RefSeq mRNA (Wang et al., 2003) sequences through dbEST or by using UNIGENE clusters (Modrek et al., 2001). Then the mRNA/ESTs themselves can be compared or they can be aligned to the genomic sequence using BLAST or specially designed software tools, such as *sim4* (http://www.hgmp.mrc.ac.uk/Registered/Webapp/sim4/) or Spidey (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/spideydoc.html), for aligning spliced cDNA sequences to genomic sequences. The alternative splicing events are inferred from the analyses of these alignments. Defining exon–intron boundaries is not a trivial task. Commonly, the alignment of two adjacent exons may overlap by 20–30 bp, and the true exon boundary may lie anywhere within this region. *Sim4* and Spidey use the score matrices of splice-site motifs to position exon boundaries. However, sometimes the inferred splice sites may deviate slightly from the correct sites, which may result in a false-positive prediction of 5′ or 3′ alternative

splice sites. Other problems that may cause errors in automated splice variant identification are coverage bias toward the 3′ end of transcripts, EST sequencing errors, contamination of EST databases with genomic or incompletely spliced mRNA sequences that may result in false predictions of intron retention events, and the presence of pseudogenes and paralogous genes in the human genome that may cause problems in EST clustering (Modrek et al., 2002; Johnson et al., 2003).

Nevertheless, bioinformatics-based studies have identified a vast number of alternatively spliced transcript variants and have estimated that 35%–59% of human genes are alternatively spliced (Modrek et al., 2002). However, the mRNA isoforms identified by automated analysis of ESTs should be taken only as a prediction and require further verification. So far, expression of the predicted splice variants has been experimentally verified only in a few studies (Brett et al., 2000; Wang et al., 2003). In these cases, RT-PCR with primers flanking the predicted alternative splice sites was used to analyze the expression of a small test set of transcripts. The low throughput of the method seems to be the bottleneck for experimental validation. Therefore, the development of high-throughput methods for the analysis of splice variants is needed to assess the specificity and sensitivity of the bioinformatics-based studies.

Recently, several large databases of alternative splicing have been developed (Huang et al., 2003; Thanaraj et al., 2004). In ProSplicer, alternative splicing forms are predicted by aligning not only ESTs and mRNAs but also protein sequences against the genomic sequence. In addition to the predicted exon–intron structures, it provides EST tissue information and links to OMIM, GO, and HUGO databases (Huang et al., 2003). A database developed by the Alternative Splicing Database (ASD) consortium contains data on AS events not only in humans, but also in mice, *Drosophila*, and other model organisms, allowing further validation of the identified splice variants by checking the preservation of the splice events across species. It also provides information about the characteristics of splice signals and provides links to GO classification and expression data (Thanaraj et al., 2004). The Extended Alternatively Spliced EST Database (EASED) project establishes a comprehensive database of alternative splice forms for nine eukaryotic organisms, including humans. The strength of this database is the provision of useful and detailed information (for humans only), for example, tissue type, developmental stage, disease notation, AS profile and classification of splice events, and the possibility of combining query parameters in order to filter out sequences of particular interest (Pospisil et al., 2004).

Exon Junction Microarrays

Johnson et al. (2003) demonstrated the use of oligonucleotide microarrays to monitor alternative exon usage. In this study, microarrays containing ~125,000 probes positioned at all known exon–exon junctions of 10,000 genes were hybridized to RNA samples derived from 52 different tissues and cell lines in order to identify tissue-specific splicing differences. The theoretical intensity of each probe was modeled as a function of the intensity of the probe response and tissue-specific expression level, and deviation between the observed and modeled intensities was used to predict an alternative splicing event. This approach proved to be useful for monitoring the expression of known splice variants and revealing novel exon-skipping events across different tissues. However, it generally is not capable of identifying novel 5′ or 3′ splice sites, of specifying the sequence of a novel splice variant, or of determining the combinatorial patterns of exon inclusion/skipping in the same transcript.

Bead-Based Fiber-Optic Microarrays

A novel approach that combines a fiber-optic microarray platform with a technique called RASL (RNA-mediated annealing, selection, and ligation) for a large-scale analysis of alternative splice variants was developed by Yeakley et al. (2002). In the RASL technique, the mRNA samples to be analyzed are annealed to multiple pairs of oligos complementary to exonic sequences flanking the presumed splice junctions. If the predicted splice variant is present, the corresponding two oligos, both carrying universal primer sites, are ligated enzymatically and amplified by PCR. The fluorescently labeled PCR products are then applied to a bead-based microarray platform that allows the quantity of each PCR product to be measured. Up to 100 splice junctions can be analyzed in a multiplexed fashion, and differences between mRNA samples in splicing efficiency and pattern can be measured.

This approach turned out to be reliable and extremely sensitive, allowing the detection of highly expressed mRNAs from less than 10 cells, which makes it particularly attractive for analyzing gene expression and splicing patterns in microdis-

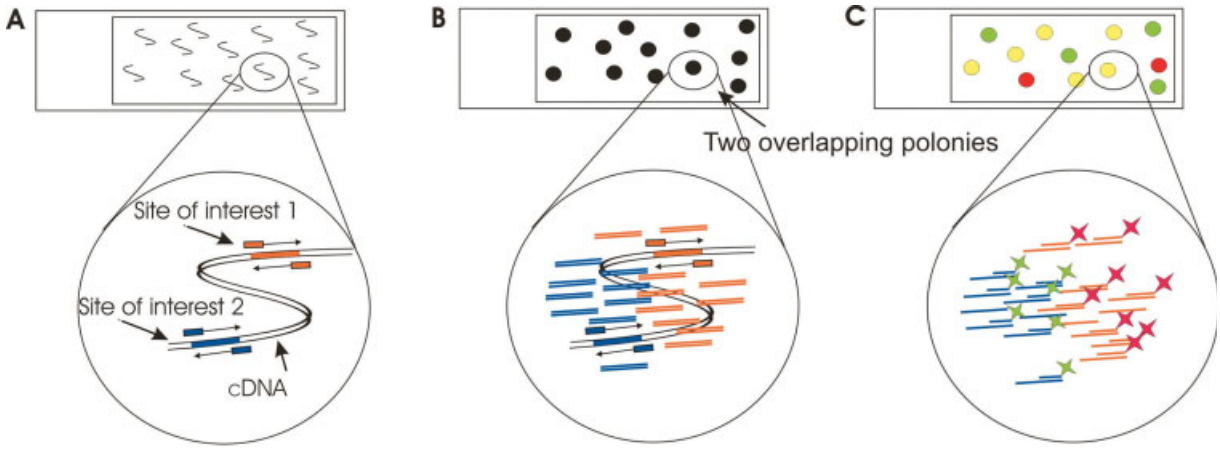

Figure 3. Polymerase colony (polony) technology. (A) Polymerization of diluted cDNA into a polyacrylamide gel containing all necessary reagents for PCR on a glass microscope slide. (B) Amplification of a single cDNA molecule with two primer pairs flanking the exons or the exon junctions of interest. Amplification products accumulate around the template, forming two overlapping polonies. One primer of each set was modified so that one strand of the amplified cDNA fragment was attached into the gel. All unattached strands were degraded and washed away after the amplification. (C) Polonies were visualized by hybridizing with exon-specific probes and scanning of slides. Green and red polonies represent transcripts with only one of the exon–exon junctions of interest; yellow polonies represent transcripts containing both exon–exon junctions of interest [adapted from Mitra et al. (2003) and Zhu et al. (2003)].

sected tissues. However, similarly to the exon junction microarrays, it requires knowledge of exon–intron structures and therefore is not applicable for searching for novel transcript variants or for the analysis of combinatorial splicing patterns.

Polymerase Colony (Polony) Technology

In the polony approach (Fig. 3), a small amount of cDNA or DNA template is polymerized into an acrylamide gel containing all the reagents necessary for PCR (Mitra et al., 2003; Zhu et al., 2003). In-gel PCR is performed on a glass microscope slide. Each template gives rise to a polony, an individual colony of amplification products. Amplification can be carried out simultaneously with several sets of primers flanking the splice sites of interest, thus generating overlapping polonies amplified from a different region of the same DNA molecule. One primer in each pair is modified in order to immobilize one strand of the amplified DNA on the acrylamide gel. This strand serves as a template for hybridization with fluorescently labeled exon-specific probes. The fluorescence is then recorded by a microarray scanner, and the images obtained are analyzed (merged) to determine the frequency of exon usage and the combinatorial patterns of exon inclusion and skipping and to quantify the expression of individual mRNA isoforms.

Currently, this is the only technique that allows multiple alternative splicing events in the same mRNA molecule to be monitored. Moreover, AS profiling can be combined with single-base extension (SBE) reactions in order to correlate SNPs with changes in the splicing pattern or efficiency (Zhu et al., 2003).

CHANGES IN PRE-mRNA SPLICING IN CANCER

A simple PubMed search with the keywords "splicing defects" or "aberrant splicing" and "cancer" found up to 100 genes whose pre-mRNA splicing is altered in various types of cancer. Both constitutive and alternative splicing may be changed, resulting in the production of novel, unnatural mRNA isoforms or in changes to the ratio or tissue specificity of natural mRNAs. Splicing defects have been observed in genes implicated in both the susceptibility and the progression of cancer. Several recent bioinformatic studies have revealed a vast number of potentially cancer-specific or cancer-associated splice variants. Wang et al. (2003) compared 11,014 mRNAs from the RefSeq database with about 3.47 million ESTs. That resulted in the identification of 26,258 alternative splice variants, of which 845 (3.2%) were significantly ($P < 0.05$) associated with cancer. The expression of 76 predicted splice variants was analyzed in paired specimens of tumor and adjacent tissue by RT-PCR, finding 72% of the expected products to be detectable and 59% to have a cancer-associated expression pattern. The analysis of 2 million ESTs in another genomewide survey of splicing aberrations identified 316 human genes that have cancer-specific splice variants (Xu et al., 2003). Hui et al. (2004) identified 383 potentially tumor-associated splice variants by aligning ESTs

with the genomic sequence of 4,322 genes. The frequencies of the identified alternative splice variants were compared between EST libraries derived from cancerous tissues with their corresponding normal counterparts and normalizing them against the expression level of each gene.

This raises a number of questions. What molecular mechanisms are responsible for changes in the recognition and usage of splice sites? Do splicing defects contribute to tumor development or progression? How frequent are splicing defects in human tumors? Is there a category of cancers in which splicing defects predominate, similar to the microsatellite instability or chromosome instability phenotype? Could the increase in splicing aberrations contribute to genetic instability? Could the protein isoforms generated as a result of splicing aberrations be recognized by a host's immune system and could they be used as targets for cancer immunotherapy?

Mechanisms of Splicing Defects

Generally, the mechanisms leading to splicing defects in cancer are poorly understood. It has been shown in individual cases that inherited or somatic mutations in *cis*-regulatory elements, as well as oncogenic signaling and variations in the composition, concentration, localization, and activity of *trans*-acting regulatory factors may result in the changes in splice-site recognition and usage.

Cis-Regulatory Elements

At least 15% of all disease-associated point mutations result in splicing defects, indicating that sequence variations in *cis*-regulatory elements could be responsible for a substantial number of missplicing events (Krawczak et al., 1992). However, whereas germ-line mutations resulting in splicing defects have been found in a number of genes implicated in a predisposition to cancer, such as *BRCA1*, *BRCA2*, *CDKN2A*, and *APC* (Liu et al., 2001; Charames et al., 2002; Agata et al., 2003; Rutter et al., 2003; Neklason et al., 2004), somatic mutations causing splicing defects rarely occur. Germ-line sequence variations in both splice sites and regulatory elements have been implicated in susceptibility to cancer. For example, an inherited nonsense mutation (Glu1694Ter) in exon 18 of the *BRCA1* gene disrupts an ESE, the binding site for SR protein SF2/ASF, causing inappropriate skipping of the constitutive exon 18 (Liu et al., 2001). Germ-line mutations in *APC* result in familial adenomatous polyposis (FAP). Mutations are scattered throughout the gene, with the precise location determining the severity of the disease (Neklason et al., 2004). Recently two mutations that disrupt splice regulatory elements have been found. The insertion of a T just beyond the conserved GT dinucleotide in the splice donor site of intron 4 results in the skipping of exon 4 and leads to an attenuated form of FAP (Neklason et al., 2004). A G-to-A substitution at the splice acceptor site of intron 7 creates a cryptic splice site, causing a single nucleotide deletion at the beginning of exon 8. A single base frameshift results in APC truncation and is associated with the classical polyposis phenotype (Charames et al., 2002).

Loss of heterozygosity and germ-line and somatic mutations in the *NF1* gene are frequently found in neurofibromas. A systematic study of somatic *NF1* mutations by Serra et al. (2001) demonstrated that most point mutations resulted in splicing defects, including exon skipping and the usage of alternative 5′ and 3′ splice sites. To our knowledge, this is the only study demonstrating that somatic mutations lead to altered splice-site recognition.

Oncogenic Signaling

CD44 is a multifunctional cell-surface glycoprotein involved in cell proliferation, differentiation, adhesion, and migration. Multiple CD44 isoforms can be generated by including and combining up to 10 variant internal exons under normal physiological conditions and during tumorigenesis. Certain CD44 splice variants, in particular those containing variant exons v5, v6, and v7, are overexpressed in various tumors and have been shown to play a role in tumor cell invasion and metastasis (Cooper et al., 1995; Naor et al., 2002; Faustino et al., 2003). The studies by Konig et al. (1998) and Matter et al. (2000) demonstrated that inclusion of exon v5 is regulated by a composite exonic splice regulator encompassing enhancer and silencer elements. The activation of the oncogenic RAS signal transduction pathway stimulates inclusion of exon v5. Because inducible splicing does not require de novo protein synthesis, the regulation of splicing is presumably carried out by a post-translational modification or by controlling the intracellular localization of splicing factors. In fact, Sam68, a member of the STAR (signal transduction and activation of RNA) protein family, was recently shown to become phosphorylated on activation of the ERK MAP-kinase pathway and was shown to be required for the RAS-induced inclusion of exon v5 (Matter et al., 2002).

Expression of *Trans*-Regulatory Elements

Several studies have demonstrated specific alterations in the expression of splicing factors in cancer. The increased phosphorylation and the elevated mRNA expression levels of Tra26, YB-1, and classical SR proteins, including SC35 and ASF/SF2, have been observed in human ovarian cancer and in mouse models of breast cancer development. Furthermore, these alterations correlate with the increased complexity of the CD44 splicing pattern (Stickeler et al., 1999; Fischer et al., 2004). Another study demonstrated that the p210 BCR/ABL7 fusion protein, generated by the Philadelphia translocation and considered responsible for initiation of chronic myelogenous leukemia, stimulates the expression of multiple genes involved in pre-mRNA splicing (e.g., SRPK1, RNA Helicase 2/Gu, hnRNPA2/B1, DDX10, and SF3B). This correlates with changes in the splicing efficiency of beta1-integrin-responsive nonreceptor tyrosin kinase (Salesse et al., 2004).

Although somatic splicing mutations are rarely found in cancers, ruling out an increased mutation rate as the principal mechanism for splicing defects, the current figures may be misleading. Generally, splicing enhancers and silencers, in particular those deep in introns, are poorly conserved, and therefore their identification and the evaluation of the impact of a nucleotide substitution on splicing efficiency relies mostly on experimental approaches on a case-by-case basis. It seems likely that many mutations currently classified as neutral, missense, or nonsense may turn out to affect the splicing process. In this context, it would be interesting to see whether tumors with mismatch repair defects also have an increased rate of splicing defects.

If the expression or activity of a splicing factor were altered in a cancer cell, it presumably would affect splicing of a subset of genes. Therefore, it should be possible to classify cancers according to their splicing defects. So far, there have been no studies done that would enable evaluation of which splicing factors have altered expression and/or activity, how frequently such alterations occur in human cancers, and what and how many genes are affected. It would also be of interest to see whether splicing defects confer a particular phenotype or have prognostic significance.

Functional Significance of Splicing Defects

At the protein level, a switch in splicing pattern can lead, for example, to an in-frame insertion or deletion of functional domains, the replacement of N or C termini, and/or truncated protein isoforms because of a frameshift. If an alternative transcript contains a premature termination codon more than 55 nucleotides upstream of the last exon–exon junction, the mRNA will most likely be degraded by the nonsense-mediated mRNA decay (NMD) pathway (Cartegni et al., 2002). AS can also generate mRNA isoforms differing in their 5′ or 3′ UTRs that may affect their stability, translation efficiency, and localization (Mignone et al., 2002). Whether splicing defects contribute to tumor development and progression or are just a consequence of the reduced fidelity of splicing machinery in cancer is still a controversial issue. Genetic instability acquired by mismatch repair defects or chromosome instability is a prerequisite for the selection of more aggressive tumor cells. Clearly, splicing machinery defects that affect a large number of genes would be lethal at the cellular level, whereas alterations in individual genes or moderate changes in the regulation of splicing might serve as additional means of increasing proteome complexity in a cancer cell (Philips et al., 2000). In some instances, aberrant splicing appears to serve as an alternative mechanism for inactivation of tumor suppressors or contributes to the gain of function of proteins promoting tumor development. Examples of aberrantly spliced genes contributing to tumor development or progression are listed in Table 1.

Inactivation of Putative Tumor Suppressors *FHIT* and *TSG101*

FHIT, a putative tumor suppressor, is a member of the histidine triad gene family encoding a diadenosine triphosphate hydrolase that is involved in purine metabolism (Barnes et al., 1996). In addition to the wild-type transcript, a vast number of different aberrant *FHIT* transcripts have been detected in various human tumors, including gastric, cervical, thyroid, and testicular germ-cell tumors (McIver et al., 2000; Lee et al., 2001, 2002; Huiping et al., 2002; Kraggerud et al., 2002; Terry et al., 2002). These transcripts are generated by exon skipping, by the use of alternative 5′ and 3′ splice sites, and by the recognition of cryptic splice sites, resulting in insertions of intronic sequences. No mutations have been found in the splice sites (McIver et al., 2000; Huiping et al., 2002), indicating that alterations in *trans*-acting regulators may be responsible for altered splice-site recognition. So far, no alternative FHIT protein isoforms have been described (Huiping et al., 2002), making it

TABLE 1. Genes Whose Splicing Pattern or Efficiency is Altered in Cancer

Gene	Splice variant	Mechanism	Functional impact	References
BRCA1	Skipped exon 18	Germ-line mutation (Glu1694Ter) disrupting ESE	Breast cancer susceptibility	Liu et al., 2001
APC	Skipped exon 4	Insertion of T beyond 5′ splice site	Attenuated FAP	Neklason et al., 2004
	1-bp deletion result in frameshift	G>A substitution created a cryptic splice site	Classical FAP	Charames et al., 2002
NF1	Exon skipping, alternative 5′ and 3′ splice sites	Various somatic mutations	Inactivation of tumor suppressor; neuroblastoma progression	Serra et al., 2001
CD44	Inclusion of exons v8–v10		Acquisition of metastatic potential	Shibuya et al., 1998
	Inclusion of exons v2–v10		Acquisition of metastatic potential	Barbour et al., 2003
	Inclusion of exons v5	Oncogene-regulated hnRNP A1 interaction to regulatory *cis*-acting splice elements	Tumor progression	Konig et al., 1998; Matter et al., 2000
FHIT	Exon skipping, alternative 5′ and 3′ splice sites	Possible alterations of *trans*-acting regulators	Inactivation of putative tumor suppressor by decreasing concentration of the functional mRNA	Huiping et al., 2002
TSG101	Exon skipping		Loss of candidate tumor-suppressor gene activity	Sun et al., 1997; McIver et al., 2000
RON (MSTIR)	Skipping of exon 11	Presence of a noncanonical splice-site consensus sequence	Acquisition of metastatic potential	Collesi et al., 1996
	Skipping of exons 5 and 6		Induction of cell scattering, transformation, and tumor growth in vivo	Zhou et al., 2003
HER2 (ERBB2)	Intron retention		Inhibition of growth factor–mediated tumor cell proliferation	Aigner et al., 2001
Fibronectin *(FN1)*	Insertion of a cryptic exon		Neovascularization/angiogenesis	Ebbinghaus et al., 2004
BIN1	Inappropriate expression of exon 12A		Abolished ability of Bin1 to inhibit malignant transformation by c-Myc or adenovirus EIA	Ge et al., 1999
MDM2	Alternative 5′ and 3′ splice sites	Aberrant splicing at direct repeat sequences	Loss of p53 binding domain. Activation of putative oncogene — promotion of p53-independent cell growth, inhibition of apoptosis	Lukas et al., 2001; Bartel et al., 2002
FAS (TNFRSF6)	Retention of intron 5		Impairment of Fas-induced apoptosis	van Doorn et al., 2002

unlikely that alternatively spliced products affect normal FHIT function in a dominant-negative manner. The splicing defects correlate with reduced FHIT protein production in most cases (McIver et al., 2000; Lee et al., 2001; Kraggerud et al., 2002). This suggests that deregulation of splicing could serve as an alternative mechanism for inactivation of the tumor suppressor by decreasing the level of functional mRNA. However, in several studies, some aberrant FHIT transcripts also were detected in normal tissues, making their implication for tumorigenesis uncertain (Gayther et al., 1997; Wang et al., 1999, 2000). This controversy presumably originated from the tissue distribution and functions of most of *FHIT* transcript variants not being characterized. It is therefore difficult to distinguish which variants represent tissue-specific alternative splicing and which represent splicing aberrations. It is also possible that aberrant transcript variants are produced because spliceosomal errors occur in both cancerous and normal tissue, but that they occur with increased frequency in cancer.

Another tumor suppressor, *TSG101*, whose expression appears to be important for the maintenance of genomic stability and cell-cycle regulation, is also a target of splicing defects and shows splicing abnormalities very similar to FHIT that are detectable in a wide range of tumors, including breast, ovary, prostate, thyroid, and cervical cancers and AML (Gayther et al., 1997; Sun et al., 1997; Klaes et al., 1999; Lin et al., 2000; McIver et al., 2000). Similarly, some transcript variants are also detectable in normal tissue, but the complexity and frequency of the aberrant splice variants is increased in tumors, showing a progressive loss of splicing fidelity during the malignant transformation (Klaes et al., 1999; Lin et al., 2000).

Aberrant *MDM2* Splicing Correlates with TP53 Status and Overall Survival

MDM2 is a putative oncogene that has been implicated in various cellular processes, including regulation of cell growth and apoptosis. MDM2 and TP53 form an autoregulatory loop in which the transcription of the *MDM2* gene is up-regulated by TP53, whereas the binding of MDM2 to TP53 inhibits transactivation by TP53 and promotes its degradation by proteasomes (Bartel et al., 2002, 2004). More than 40 *MDM2* transcript variants have been identified, but the pattern of alternative splicing and the number of transcripts differ between normal and cancerous tissues. The encoded proteins are likely to be functionally different. Thus, for example, MDM2-b, one of the splice forms lacking a TP53-binding domain has been shown to promote TP53-independent cell growth, to inhibit apoptosis, and to up-regulate the RelA subunit of NFkappaB (Steinman et al., 2004). Some transcripts generated by the use of cryptic splice sites and the skipping of exons were detectable exclusively in invasive breast cancer, not in normal breast tissue. No mutations that could create these cryptic splice sites were found in these tumors, but sequence analysis demonstrated that splicing occurred between direct repeat sequences. Moreover, the expression of the aberrant transcripts was correlated with *TP53* overexpression, *TP53* mutations, and shorter overall survival of breast cancer patients, suggesting that aberrant splice forms may play a significant role in tumorigenesis and might serve as a prognostic marker for breast cancer (Lukas et al., 2001).

Generation of Constitutively Active RON Tyrosine Kinase Isoforms

Another example of how alternative splicing contributes to tumor growth and the acquisition of an invasive phenotype is RON (MST1R according to the HUGO nomenclature). RON is a tyrosine kinase receptor for macrophage-stimulating protein (MSP). Wild-type RON is a heterodimeric protein composed of α and β subunits derived from proteolytic cleavage of a common precursor. The binding of MSP to RON triggers an intracellular signaling cascade eliciting cell dissociation, motility, and invasiveness (Comoglio et al., 1996). Three alternatively spliced RON transcripts, encoding 165-, 160-, and 155-kDa RON isoforms, have been identified. RONΔ165 was found in a gastric cancer cell line (Collesi et al., 1996), whereas all three isoforms were detected in primary colorectal carcinomas (Zhou et al., 2003). All three RON variants are constitutively active but differ in their biochemical and biological properties (Zhou et al., 2003). RONΔ165 is generated by skipping a 147-bp exon, resulting in a 49-amino-acid deletion from the extracellular domain, which prevents the precursor from undergoing proteolytic cleavage. The unprocessed precursor is retained in the cytoplasm, and the tyrosine kinase is constitutively activated by oligomerization and autophosphorylation. Cells expressing RONΔ165 acquired invasive properties in vitro (Collesi et al., 1996). RONΔ155 also is not cleaved and is present as a single-chain precursor. RONΔ160 is produced by skipping exons 5 and 6 (109 amino acids from the extracellular domain),

but this does not interfere with the cleavage of the precursor, and the heterodimeric 160-kDa protein is located on the cell surface and binds MSP. Both the 160- and 155-kDa isoforms induce cell scattering, transformation, and tumor growth in vivo (Zhou et al., 2003).

IMMUNOLOGICAL RECOGNITION OF ABERRANTLY SPLICED GENE PRODUCTS IN CANCER

Recently, it has been shown that frameshift-mutation-derived peptides are recognized by tumor-infiltrating lymphocytes in MSI+ colon cancer patients, and it has been suggested that these are very attractive targets for cancer vaccines (Saeterdal et al., 2001; Ishikawa et al., 2003). Similarly, peptides derived form aberrantly spliced transcripts could be recognized by the immune system and serve as targets for cancer immunotherapy. Novel exon–exon junctions generated by exon skipping, intron retention (in-frame and out-of-frame translations), and the inclusion of cryptic exons may result in structural alterations of a protein and represent novel T- and B-cell epitopes. In fact, several proteins encoded by alternatively or aberrantly spliced genes have been shown to induce an immune response in cancer patients (Table 2).

Peptides Derived from Aberrantly Spliced Transcripts as CTL Targets

Yannik Guilloux et al. (1996) identified a peptide derived from an intronic sequence of the N-acetylglucosaminyltransferase V (*GnT-V/MGAT5*) gene as a novel tumor antigen, NA17-A, that is recognized by cytotoxic T lymphocytes (CTLs) on HLA-A2 melanomas. The mRNA encoding the NA17-A antigen is generated by the use of an alternative promoter in an intron and is translated in a different reading frame than *GnT-V*. The aberrant transcript was expressed at a relatively high level in 50% of HLA-A2 melanomas, whereas only trace levels were detectable in normal tissues (Guilloux et al., 1996). Peptide-pulsed and apoptotic body-loaded dendritic cells stimulate NA17-A-specific CTL activation, demonstrating that NA17-A could be a useful target for cancer immunotherapy (Labarriere et al., 1998, 2002).

The tyrosinase-related protein (TRP) 2-INT2 isoform represents another case of intron retention that results in the generation of a melanoma-restricted T-cell epitope. *TRP2* itself encodes melanoma differentiation antigen, which has its own antigenic epitope expressed in melanoma and melanocytes (Wang et al., 1996). The TRP2-INT2 isoform contains exons 1–4, with retention of intron 2 and a part of intron 4. The translation of intron 2 in the same reading frame of the fully spliced TRP2 mRNA introduces a stop codon at the 3′ end of the intron. In contrast to TRP2, this isoform is highly expressed only in melanoma cells, not in normal melanocytes, suggesting that the splicing pattern of TRP2 is altered in melanoma. The peptide encoded by the intronic sequence is recognized by CTLs when exposed on HLA-A*68011 (Lupetti et al., 1998). Recently another TRP2 isoform, TRP2-6b, generated by the insertion of two cryptic exons (6b and 6c) between exon 6 and 7, was identified. It is translated in-frame with the wt TRP2 and contains an insertion of 33 amino acids. Two peptides encoded by exon 6b were recognized by melanoma-specific HLA-A2 restricted tumor-infiltrated lymphocytes (Khong et al., 2002).

Autoantibody Responses Against Alternatively Spliced Gene Products

The SEREX (serological identification of antigens by recombinant expression cloning) approach is based on the screening of cDNA expression libraries with cancer patients' sera, allowing a systematic search for genes whose products have elicited autoantibody production in cancer patients (Sahin et al., 1995, 2001). This technique has been applied to various types of tumors, resulting in the identification of approximately 1,450 potential tumor antigens. So far, the molecular mechanisms underlying the immunogenicity of the majority of SEREX-identified antigens are unknown, but it is assumed that the antibody responses may be associated with a cancer-associated expression pattern (cancer-testis and differentiation antigens), mutations, and overexpression (Türeci et al., 1999; Pfreundschuh, 2000). A large number of these genes have multiple transcript variants generated by alternative promoter and splice-site usage, raising the possibility that alterations in splicing might be associated with their immunogenicity. In fact, alterations in splicing pattern or efficiency and the use of alternative promoters in cancer have been demonstrated for several genes, including *ELAV* (Hu antigen D), *TACC1*, *CML66*, *CTAGE*, and *XAGE1* (Behrends et al., 2002; Line et al., 2002a; Usener et al., 2003; Ali Eldib et al., 2004; Yan et al., 2004).

We previously identified a microtubule-associated protein, TACC1, as a potential tumor antigen by applying the SEREX technique to gastric cancer (Line et al., 2002a, 2002b). The serum-reactive cDNA clone represented a novel splice variant of

TABLE 2. Splice Variants Inducing Cellular or Humoral Immune Responses in Cancer Patients

Gene	Splicing pattern	Tumor type	Expression pattern	T- or B-cell response	Epitope	References
XAGE1 (GAGED2)	4 transcript variants: 1a, 1b, 1c, 1d	Lung cancer	Cancer — testes	Autoantibody responses to XAGE-1b	ND	Ali Eldib et al., 2004
				Autoantibody responses to XAGE-1c (L552S)	N-terminal peptide unique to 1c isoform	Wang et al., 2001b
TACC1	10 transcript variants (*TACC1-A–TACC1-J*)	Gastric cancer	*TACC1-F*: normal brain and gastric cancer; *TACC1-D*: gastric cancer, trace amounts in normal brain and stomach	Autoantibody responses to TACC1	ND	Line et al., 2002a, 2002b
GnT-V *(MGAT5)*	2 transcript variants generated by usage of alternative promoter	Malignant melanoma	*NA17-A* encoding isoform expressed at high levels in melanomas, trace levels detectable in normal tissues	CTL responses to antigen NA17-A	Peptide generated from retained intron	Guilloux et al., 1996
ELAV (Hu antigen D)	4 transcript variants: *HuD1–HuD4*	Childhood neuroblastoma	*HUD1* expression restricted to CNS and neuroectodermal tumors	Autoantibody responses to HUD1	ND	Behrends et al., 2002
CML66	2 transcript variants: *CML66-L* and *–S* *TRP 2-INT2*	CML	*CML66-L*: cancer—testes; *CML66-S*: testes, barely in cancer	Autoantibody responses to CML66-L	5' region of protein isoform CML66-L	Yan et al., 2004
TRP2 (TYRP1)	*TRP2-INT2* (retention of intron 2)	Malignant melanoma	Melanoma—specific	CTL responses to TRP-2-INT2 and TRP-2-6b	Peptides encoded by intron 2	Wang et al., 1996; Khong et al., 2002
	TRP-2-6b (insertion of novel exons 6b and 6c)		?		2 peptides encoded by exon 6b	Khong et al., 2002
CTAGE-1	3 transcript variants: *cTAGE-1, -1B, -1C*	Cutaneous T cell lymphoma	*CTAGE*-1 and -1B: cancer—testes	Autoantibody responses to cTAGE-1	Identified for cTAGE-1	Usener et al., 2003
TAG	4 transcript variants: TAG-1, *TAG-2a, TAG-2b, TAG-2c*	Melanoma	Cancer—testes	CTL responses to TAG-derived peptide	Peptide encoded by junction of the first two exons	Hogan et al., 2004
ING1	4 transcript variants: A, B, C, D	Breast cancer	A — ubiquitous B — differential, some tumors; C — testes, weakly in brain	Autoantibody responses to ING1	ND	Jager et al., 1999

ND, not determined.

TACC1 generated by insertion of a 36-bp exon. Analysis of the *TACC1* mRNA structure and expression pattern revealed 10 novel transcript variants and demonstrated that the regulation of alternative splicing of *TACC1* pre-mRNA is altered in gastric cancer, resulting in overexpression of two splice variants, *TACC1-F* and *-D*. The two transcript variants share the same 36-bp exon, which is not included in any other transcript. *TACC1-F* was predominantly expressed in normal brain and at comparable levels in 50% of gastric cancer specimens, which was 50- to 1,000-fold greater than the expression in paired adjacent tissue and in other normal tissue. *TACC1-D* was expressed in 52% of gastric tumors, but only trace levels were detectable in adjacent gastric mucosa and normal brain (unpublished data). It seems likely that the production of *TACC1-F* and *-D* isoforms in cancerous tissues may underlie the immunogenicity of TACC1. However, which region of the protein is recognized by autoantibodies and whether the isoform-specific peptides are capable of inducing CTL responses remain to be determined.

XAGE-1 (G antigen, family D) belongs to the family of cancer-testis antigens. Four transcript variants are generated by alternative splicing and alternate transcription initiation sites. Recently, two of these isoforms—XAGE-1b and 1c, were detected by SEREX analysis of lung cancer (Ali Eldib et al., 2004). In this study, XAGE-1b was shown to be the dominant transcript in testis and lung cancer, and it was recognized by sera from 8 of 32 cancer patients but not by healthy controls. In a study by Wang et al. (2001b), another XAGE-1 isoform, L552S (corresponding to XAGE-1c), was shown to be overexpressed and immunogenic in lung cancer. Epitope mapping showed that the N-terminal region, which is unique in the XAGE-1c isoform, was recognized by autoantibodies in at least one patient.

Six antigens encoded by novel splice variants were identified by SEREX in chronic lymphocytic leukemia (CLL). One of the clones, KW-4, encodes a Kruppel-like zinc finger protein, ZNF286, missing the KRAB domain. Full-length ZNF286 is ubiquitously expressed, whereas the identified splice variant was not detected in 8 normal tissue samples but was expressed in all CLL cases tested (Krackhardt et al., 2002). However, it is not known whether the immune response is directed against the splice-variant-specific epitope.

Recently, six transcript variants of a microtubule-associated protein, kinectin, were found to induce antibody responses in patients with hepatocellular carcinoma (Wang et al., 2004). One of the variants showed cancer-associated overexpression, suggesting a putative association of the altered splicing efficiency with the immunogenicity.

We have described here how splicing defects may contribute to the generation of novel epitopes. However, it is possible that the alterations in splicing also could change the antigen repertoire of tumor cells by removing immunodominant epitopes, contributing to the immune escape of tumors. In fact, the differential splicing of RNA encoding a CTL epitope of EBV has been shown to reduce the functional presentation of the immunodominant EBV epitope (Kienzle et al., 2000).

In several instances described here (TRP2-INT2, HER2, FAS, etc.), missplicing events introduced premature termination codons. Thus, in theory, these transcripts should be targeted for degradation by the NMD pathway, a conserved proofreading mechanism in eukaryotic cells that protects them from the potentially detrimental effects of truncated proteins (Byers, 2002). Nevertheless, we have summarized the evidence that aberrant transcripts can be translated into functional proteins, generating novel cancer-specific antigenic epitopes recognized by both the cellular and humoral responses of the host's immune system. This raises an intriguing possibility: that the NMD mechanism by itself might be disturbed in cancers, resulting in the accumulation of aberrant transcripts that escape from NMD.

We have attempted to provide insight into the mechanisms and the significance of pre-mRNA splicing alterations in cancer. We have only mentioned a small number of genes whose splicing is altered in cancer. These alterations result either in reduction of the number of normal transcripts, thus contributing to inactivation of tumor suppressors, or in the production of novel protein isoforms conferring oncogenic or metastatic potential to the tumor cells. In some cases, the proteins generated from the aberrant transcripts can be recognized by the host's immune system, revealing their relevance as immunotherapeutic targets. However, a line of evidence suggests that splicing defects are an essential feature of cancer. New tools allowing systematic large-scale analyses of alternative splicing efficiency and patterns are emerging. Hopefully, this will lead to a better understanding of the nature and the role of splicing defects in cancer. We anticipate that the determination of splicing patterns in cancer will provide novel diagnostic and prognostic markers and may serve as a source

for the identification of novel targets for the immunotherapy of cancer. Furthermore, the splice variants that functionally contribute to tumor development or progression may be targeted by various molecular therapies that appear to be close to clinical application, which include targeting by small molecules of aberrant protein isoforms or splicing factors, inhibition or activation of specific splicing events by the use of various oligonucleotide-mediated therapies, and RNA-based reprogramming of alternative splicing (for a detailed review, see Garcia-Blanco et al., 2004).

REFERENCES

Agata S, De Nicolo A, Chieco-Bianchi L, D'Andrea E, Menin C, Montagna M. 2003. The BRCA2 sequence variant IVS19+1G→A A leads to an aberrant transcript lacking exon 19. Cancer Genet Cytogenet 141:175–176.

Aigner A, Juhl H, Malerczyk C, Tkybusch A, Benz CC, Czubayko F. 2001. Expression of a truncated 100 kDa HER2 splice variant acts as an endogenous inhibitor of tumour cell proliferation. Oncogene 20:2101–2111.

Ali Eldib AM, Ono T, Shimono M, Kaneko M, Nakagawa K, Tanaka R, Noguchi Y, Nakayama E. 2004. Immunoscreening of a cDNA library from a lung cancer cell line using autologous patient serum: identification of XAGE-1b as a dominant antigen and its immunogenicity in lung adenocarcinoma. Int J Cancer 108:558–563.

Barbour AP, Reeder JA, Walsh MD, Fawcett J, Antalis TM, Gotley DC. 2003. Expression of the CD44v2-10 isoform confers a metastatic phenotype: importance of the heparan sulfate attachment site CD44v3. Cancer Res 63:887–892.

Barnes LD, Garrison PN, Siprashvili Z, Guranowski A, Robinson AK, Ingram SW, Croce CM, Ohta M, Huebner K. 1996. Fhit, a putative tumor suppressor in humans, is a dinucleoside 5′,5″-P1,P3-triphosphate hydrolase. Biochemistry 35:11529–11535.

Bartel F, Taubert H, Harris LC. 2002. Alternative and aberrant splicing of MDM2 mRNA in human cancer. Cancer Cell 2:9–15.

Bartel F, Harris LC, Wurl P, Taubert H. 2004. MDM2 and its splice variant messenger RNAs: expression in tumors and down-regulation using antisense oligonucleotides. Mol Cancer Res 2:29–35.

Behrends U, Jandl T, Golbeck A, Lechner B, Muller-Weihrich S, Schmid I, Till H, Berthold F, Voltz R, Mautner JM. 2002. Novel products of the HUD, HUC, NNP-1 and alpha-internexin genes identified by autologous antibody screening of a pediatric neuroblastoma library. Int J Cancer 100:669–677.

Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72:291–336.

Brett D, Hanke J, Lehmann G, Haase S, Delbruck S, Krueger S, Reich J, Bork P. 2000. EST comparison indicates 38% of human mRNAs contain possible alternative splice forms. FEBS Lett 474:83–86.

Burset M, Seledtsov IA, Solovyev VV. 2000. Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Res 28:4364–4375.

Byers PH. 2002. Killing the messenger: new insights into nonsense-mediated mRNA decay. J Clin Invest 109:3–6.

Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285–298.

Charames GS, Cheng H, Gilpin CA, Hunter AG, Berk T, Bapat B. 2002. A novel aberrant splice site mutation in the APC gene. J Med Genet 39:754–757.

Collesi C, Santoro MM, Gaudino G, Comoglio PM. 1996. A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype. Mol Cell Biol 16: 5518–5526.

Comoglio PM, Boccaccio C. 1996. The HGF receptor family: unconventional signal transducers for invasive cell growth. Genes Cells 1:347–354.

Cooper DL, Dougherty GJ. 1995. To metastasize or not? Selection of CD44 splice sites. Nat Med 1:635–637.

Dredge BK, Darnell RB. 2003. Nova regulates GABA(A) receptor gamma2 alternative splicing via a distal downstream UCAU-rich intronic splicing enhancer. Mol Cell Biol 23:4687–4700.

Dredge BK, Polydorides AD, Darnell RB. 2001. The splice of life: alternative splicing and neurological disease. Nat Rev Neurosci 2:43–50.

Ebbinghaus C, Scheuermann J, Neri D, Elia G. 2004. Diagnostic and therapeutic applications of recombinant antibodies: targeting the extra-domain B of fibronectin, a marker of tumor angiogenesis. Curr Pharm Des 10:1537–1549.

Fairbrother WG, Yeh RF, Sharp PA, Burge CB. 2002. Predictive identification of exonic splicing enhancers in human genes. Science 297:1007–1013.

Faustino NA, Cooper TA. 2003. Pre-mRNA splicing and human disease. Genes Dev 17:419–437.

Fischer DC, Noack K, Runnebaum IB, Watermann DO, Kieback DG, Stamm S, Stickeler E. 2004. Expression of splicing factors in human ovarian cancer. Oncol Rep 11:1085–1090.

Garcia-Blanco MA, Baraniak AP, Lasda EL. 2004. Alternative splicing in disease and therapy. Nat Biotechnol 22:535–546.

Gayther SA, Barski P, Batley SJ, Li L, de Foy KA, Cohen SN, Ponder BA, Caldas C. 1997. 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 15:2119–2126.

Ge K, DuHadaway J, Du W, Herlyn M, Rodeck U, Prendergast GC. 1999. Mechanism for elimination of a tumor suppressor: aberrant splicing of a brain-specific exon causes loss of function of Bin1 in melanoma. Proc Natl Acad Sci USA 96:9689–9694.

Guilloux Y, Lucas S, Brichard VG, Van Pel A, Viret C, De Plaen E, Brasseur F, Lethe B, Jotereau F, Boon T. 1996. A peptide recognized by human cytolytic T lymphocytes on HLA-A2 melanomas is encoded by an intron sequence of the N-acetylglucosaminyltransferase V gene. J Exp Med 183:1173–1183.

Hogan KT, Coppola MA, Gatlin CL, Thompson LW, Shabanowitz J, Hunt DF, Engelhard VH, Ross MM, Slingluff CL, Jr. 2004. Identification of novel and widely expressed cancer/testis gene isoforms that elicit spontaneous cytotoxic T-lymphocyte reactivity to melanoma. Cancer Res 64:1157–1163.

Huang HD, Horng JT, Lee CC, Liu BJ. 2003. ProSplicer: a database of putative alternative splicing information derived from protein, mRNA and expressed sequence tag sequence data. Genome Biol 4:R29.

Hui L, Zhang X, Wu X, Lin Z, Wang Q, Li Y, Hu G. 2004. Identification of alternatively spliced mRNA variants related to cancers by genome-wide ESTs alignment. Oncogene 23:3013–3023.

Huiping C, Kristjansdottir S, Bergthorsson JT, Jonasson JG, Magnusson J, Egilsson V, Ingvarsson S. 2002. High frequency of LOH, MSI and abnormal expression of FHIT in gastric cancer. Eur J Cancer 38:728–735.

Ishikawa T, Fujita T, Suzuki Y, Okabe S, Yuasa Y, Iwai T, Kawakami Y. 2003. Tumor-specific immunological recognition of frameshift-mutated peptides in colon cancer with microsatellite instability. Cancer Res 63:5564–5572.

Jager D, Stockert E, Scanlan MJ, Gure AO, Jager E, Knuth A, Old LJ, Chen YT. 1999. Cancer-testis antigens and ING1 tumor suppressor gene product are breast cancer antigens: characterization of tissue-specific ING1 transcripts and a homologue gene. Cancer Res 59:6197–6204.

Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD. 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science 302:2141–2144.

Kan Z, Rouchka EC, Gish WR, States DJ. 2001. Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. Genome Res 11:889–900.

Khong HT, Rosenberg SA. 2002. Pre-existing immunity to tyrosinase-related protein (TRP)-2, a new TRP-2 isoform, and the NY-ESO-1 melanoma antigen in a patient with a dramatic response to immunotherapy. J Immunol 168:951–956.

Kienzle N, Buck M, Silins SL, Burrows SR, Moss DJ, Winterhalter A, Brooks A, Khanna R. 2000. Differential splicing of antigen-encoding RNA reduces endogenous epitope presentation that regulates the expansion and cytotoxicity of T cells. J Immunol 165:1840–1846.

Klaes R, Kloor M, Willeke F, Melsheimer P, von Knebel DM, Ridder R. 1999. Significant increase of a specific variant TSG101 transcript during the progression of cervical neoplasia. Eur J Cancer 35:733–737.

Konig H, Ponta H, Herrlich P. 1998. Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. EMBO J 17:2904–2913.

Krackhardt AM, Witzens M, Harig S, Hodi FS, Zauls AJ, Chessia M, Barrett P, Gribben JG. 2002. Identification of tumor-associated antigens in chronic lymphocytic leukemia by SEREX. Blood 100:2123–2131.

Kraggerud SM, Aman P, Holm R, Stenwig AE, Fossa SD, Nesland JM, Lothe RA. 2002. Alterations of the fragile histidine triad gene, FHIT, and its encoded products contribute to testicular germ cell tumorigenesis. Cancer Res 62:512–517.

Krawczak M, Reiss J, Cooper DN. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet 90:41–54.

Labarriere N, Pandolfino MC, Raingeard D, Le Guiner S, Diez E, Le Drean E, Dreno B, Jotereau F. 1998. Frequency and relative fraction of tumor antigen-specific T cells among lymphocytes from melanoma-invaded lymph nodes. Int J Cancer 78:209–215.

Labarriere N, Bretaudeau L, Gervois N, Bodinier M, Bougras G, Diez E, Lang F, Gregoire M, Jotereau F. 2002. Apoptotic body-loaded dendritic cells efficiently cross-prime cytotoxic T lymphocytes specific for NA17-A antigen but not for Melan-A/MART-1 antigen. Int J Cancer 101:280–286.

Lander ES, Linton LM, Birren B, Nusbaum C et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860–921.

Lee SH, Kim CJ, Park HK, Koh JW, Cho MH, Baek MJ, Lee MS. 2001. Characterization of aberrant FHIT transcripts in gastric adenocarcinomas. Exp Mol Med 33:124–130.

Lee SH, Kim HY, Kim TJ, Park HK, Kim WH, Woo KM, Cho MH. 2002. Aberrant splicing of FHIT transcripts in human gastric cancer cell lines. Res Commun Mol Pathol Pharmacol 112:39–49.

Lin SF, Lin PM, Liu TC, Chang JG, Sue YC, Chen TP. 2000. Clinical implications of aberrant TSG101 transcripts in acute myeloblastic leukemia. Leuk Lymphoma 36:463–466.

Line A, Slucka Z, Stengrevics A, Li G, Rees RC. 2002a. Altered splicing pattern of TACC1 mRNA in gastric cancer. Cancer Genet Cytogenet 139:78–83.

Line A, Stengrevics A, Slucka Z, Li G, Jankevics E, Rees RC. 2002b. Serological identification and expression analysis of gastric cancer-associated genes. Br J Cancer 86:1824–1830.

Liu HX, Zhang M, Krainer AR. 1998. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev 12:1998–2012.

Liu HX, Chew SL, Cartegni L, Zhang MQ, Krainer AR. 2000. Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. Mol Cell Biol 20:1063–1071.

Liu HX, Cartegni L, Zhang MQ, Krainer AR. 2001. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nat Genet 27:55–58.

Lukas J, Gao DQ, Keshmeshian M, Wen WH, Tsao-Wei D, Rosenberg S, Press MF. 2001. Alternative and aberrant messenger RNA splicing of the mdm2 oncogene in invasive breast cancer. Cancer Res 61:3212–3219.

Lupetti R, Pisarra P, Verrecchia A, Farina C, Nicolini G, Anichini A, Bordignon C, Sensi M, Parmiani G, Traversari C. 1998. Translation of a retained intron in tyrosinase-related protein (TRP) 2 mRNA generates a new cytotoxic T lymphocyte (CTL)-defined and shared human melanoma antigen not expressed in normal cells of the melanocytic lineage. J Exp Med 188:1005–1016.

Maniatis T, Tasic B. 2002. Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature 418:236–243.

Matter N, Marx M, Weg-Remers S, Ponta H, Herrlich P, Konig H. 2000. Heterogeneous ribonucleoprotein A1 is part of an exon-specific splice-silencing complex controlled by oncogenic signaling pathways. J Biol Chem 275:35353–35360.

Matter N, Herrlich P, Konig H. 2002. Signal-dependent regulation of splicing via phosphorylation of Sam68. Nature 420:691–695.

McIver B, Grebe SK, Wang L, Hay ID, Yokomizo A, Liu W, Goellner JR, Grant CS, Smith DI, Eberhardt NL. 2000. FHIT and TSG101 in thyroid tumours: aberrant transcripts reflect rare abnormal RNA processing events of uncertain pathogenetic or clinical significance. Clin Endocrinol (Oxf) 52:749–757.

Mignone F, Gissi C, Liuni S, Pesole G. 2002. Untranslated regions of mRNAs. Genome Biol 3:REVIEWS0004-

Mitra RD, Butty VL, Shendure J, Williams BR, Housman DE, Church GM. 2003. Digital genotyping and haplotyping with polymerase colonies. Proc Natl Acad Sci USA 100:5926–5931.

Modrek B, Lee C. 2002. A genomic view of alternative splicing. Nat Genet 30:13–19.

Modrek B, Resch A, Grasso C, Lee C. 2001. Genome-wide detection of alternative splicing in expressed sequences of human genes. Nucleic Acids Res 29:2850–2859.

Naor D, Nedvetzki S, Golan I, Melnik L, Faitelson Y. 2002. CD44 in cancer. Crit Rev Clin Lab Sci 39:527–579.

Neklason DW, Solomon CH, Dalton AL, Kuwada SK, Burt RW. 2004. Intron 4 mutation in APC gene results in splice defect and attenuated FAP phenotype. Fam Cancer 3:35–40.

Pagani F, Baralle FE. 2004. Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet 5:389–396.

Patel AA, Steitz JA. 2003. Splicing double: insights from the second spliceosome. Nat Rev Mol Cell Biol 4:960–970.

Patel NA, Apostolatos HS, Mebert K, Chalfant CE, Watson JE, Pillay TS, Sparks J, Cooper DR. 2004. Insulin regulates protein kinase CbetaII alternative splicing in multiple target tissues: development of a hormonally responsive heterologous minigene. Mol Endocrinol 18:899–911.

Pfreundschuh M. 2000. Exploitation of the B cell repertoire for the identification of human tumor antigens. Cancer Chemother Pharmacol 46Suppl:S3–S7.

Philips AV, Cooper TA. 2000. RNA processing and human disease. Cell Mol Life Sci 57:235–249.

Pospisil H, Herrmann A, Bortfeldt RH, Reich JG. 2004. EASED: Extended Alternatively Spliced EST Database. Nucleic Acids Res 32:D70–D74.

Rutter JL, Goldstein AM, Davila MR, Tucker MA, Struewing JP. 2003. CDKN2A point mutations D153spl(c.457G>T) and IVS2+1G>T result in aberrant splice products affecting both p16INK4a and p14ARF. Oncogene 22:4444–4448.

Saeterdal I, Bjorheim J, Lislerud K, Gjertsen MK, Bukholm IK, Olsen OC, Nesland JM, Eriksen JA, Moller M, Lindblom A, Gaudernack G. 2001. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. Proc Natl Acad Sci USA 98:13255–13260.

Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, Stenner F, Luo G, Schobert I, Pfreundschuh M. 1995. Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci USA 92: 11810–11813.

Sahin U, Li G, Tureci O, Pfreundschuh M. 2001. Recognition of human tumors: SEREX expression cloning to identify tumour antigens. In: Robins RA, Rees RC (eds). Cancer Immunology. Dordrecht/Boston/London: Kluwer Academic Publishers. p 45–57.

Salesse S, Dylla SJ, Verfaillie CM. 2004. p210BCR/ABL-induced alteration of pre-mRNA splicing in primary human CD34+ hematopoietic progenitor cells. Leukemia 18:727–733.

Scotet E, Houssaint E. 1998. Exon III splicing switch of fibroblast growth factor (FGF) receptor-2 and -3 can be induced by FGF-1 or FGF-2. Oncogene 17:67–76.

Serra E, Ars E, Ravella A, Sanchez A, Puig S, Rosenbaum T, Estivill X, Lazaro C. 2001. Somatic NF1 mutational spectrum in benign neurofibromas: mRNA splice defects are common among point mutations. Hum Genet 108:416–429.

Shen H, Kan JL, Green MR. 2004. Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol Cell 13:367–376.

Shibuya Y, Okabayashi T, Oda K, Tanaka N. 1998. Ratio of CD44 epithelial variant to CD44 hematopoietic variant is a useful prognostic indicator in gastric and colorectal carcinoma. Jpn J Clin Oncol 28:609–614.

Stamm S. 2002. Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. Hum Mol Genet 11:2409–2416.

Steinman HA, Burstein E, Lengner C, Gosselin J, Pihan G, Duckett CS, Jones SN. 2004. An alternative splice form of Mdm2 induces p53-independent cell growth and tumorigenesis. J Biol Chem 279:4877–4886.

Stickeler E, Kittrell F, Medina D, Berget SM. 1999. Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. Oncogene 18:3574–3582.

Sun Z, Pan J, Bubley G, Balk SP. 1997. Frequent abnormalities of TSG101 transcripts in human prostate cancer. Oncogene 15:3121–3125.

ten Dam GB, Zilch CF, Wallace D, Wieringa B, Beverley PC, Poels LG, Screaton GR. 2000. Regulation of alternative splicing of CD45 by antagonistic effects of SR protein splicing factors. J Immunol 164:5287–5295.

Terry G, Ho L, Londesborough P, Cuzick J. 2002. Abnormal FHIT expression profiles in cervical intraepithelial neoplastic (CIN) lesions. Br J Cancer 86:376–381.

Thanaraj TA, Stamm S, Clark F, Riethoven JJ, Le T, V, Muilu J. 2004. ASD: the Alternative Splicing Database. Nucleic Acids Res 32:D64–D69.

Tureci O, Sahin U, Zwick C, Neumann F, Pfreundschuh M. 1999. Exploitation of the antibody repertoire of cancer patients for the identification of human tumor antigens. Hybridoma 18: 23–28.

Usener D, Schadendorf D, Koch J, Dubel S, Eichmuller S. 2003. cTAGE: a cutaneous T cell lymphoma associated antigen family with tumor-specific splicing. J Invest Dermatol 121:198–206.

Van Buskirk C, Schupbach T. 2002. Half pint regulates alternative splice site selection in Drosophila. Dev Cell 2:343–353.

van Doorn R, Dijkman R, Vermeer MH, Starink TM, Willemze R, Tensen CP. 2002. A novel splice variant of the Fas gene in patients with cutaneous T-cell lymphoma. Cancer Res 62:5389–5392.

Wang RF, Appella E, Kawakami Y, Kang X, Rosenberg SA. 1996. Identification of TRP-2 as a human tumor antigen recognized by cytotoxic T lymphocytes. J Exp Med 184:2207–2216.

Wang NM, Chang JG. 1999. Are aberrant transcripts of FHIT, TSG101, and PTEN/MMAC1 oncogenesis related? Int J Mol Med 3:491–495.

Wang NM, Chang JG, Liu TC, Lin SF, Peng CT, Tsai FJ, Tsai CH. 2000. Aberrant transcripts of FHIT, TSG101 and PTEN/MMAC1 genes in normal peripheral mononuclear cells. Int J Oncol 16:75–80.

Wang HY, Xu X, Ding JH, Bermingham JR, Jr., Fu XD. 2001a. SC35 plays a role in T cell development and alternative splicing of CD45. Mol Cell 7:331–342.

Wang T, Fan L, Watanabe Y, McNeill P, Fanger GR, Persing DH, Reed SG. 2001b. L552S, an alternatively spliced isoform of XAGE-1, is over-expressed in lung adenocarcinoma. Oncogene 20:7699–7709.

Wang Z, Lo HS, Yang H, Gere S, Hu Y, Buetow KH, Lee MP. 2003. Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer. Cancer Res 63:655–657.

Wang HC, Su YR, Han KJ, Pang XW, Peng JR, Liang B, Wang S, Chen WF. 2004. Multiple variants and a differential splicing pattern of kinectin in human hepatocellular carcinoma. Biochem Cell Biol 82:321–327.

Xie H, Zhu WY, Wasserman A, Grebinskiy V, Olson A, Mintz L. 2002. Computational analysis of alternative splicing using EST tissue information. Genomics 80:326–330.

Xu Q, Modrek B, Lee C. 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res 30:3754–3766.

Xu Q, Lee C. 2003. Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences. Nucleic Acids Res 31:5635–5643.

Yan Y, Phan L, Yang F, Talpaz M et al. 2004. A novel mechanism of alternative promoter and splicing regulates the epitope generation of tumor antigen CML66-L. J Immunol 172:651–660.

Yeakley JM, Fan JB, Doucet D, Luo L, Wickham E, Ye Z, Chee MS, Fu XD. 2002. Profiling alternative splicing on fiber-optic arrays. Nat Biotechnol 20:353–358.

Zhou YQ, He C, Chen YQ, Wang D, Wang MH. 2003. Altered expression of the RON receptor tyrosine kinase in primary human colorectal adenocarcinomas: generation of different splicing RON variants and their oncogenic potential. Oncogene 22:186–197.

Zhu J, Mayeda A, Krainer AR. 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol Cell 8:1351–1361.

Zhu J, Shendure J, Mitra RD, Church GM. 2003. Single molecule profiling of alternative pre-mRNA splicing. Science 301:836–838.