Alternative Splicing and Its Impact as a Cancer Diagnostic Marker

Yun-Ji Kim, Heui-Soo Kim*

Department of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea

Most genes are processed by alternative splicing for gene expression, resulting in the complexity of the transcriptome in eukaryotes. It allows a limited number of genes to encode various proteins with intricate functions. Alternative splicing is regulated by genetic mutations in cis-regulatory factors and epigenetic events. Furthermore, splicing events occur differently according to cell type, developmental stage, and various diseases, including cancer. Genome instability and flexible proteomes by alternative splicing could affect cancer cells to grow and survive, leading to metastasis. Cancer cells that are transformed by aberrant and uncontrolled mechanisms could produce alternative splicing to maintain and spread them continuously. Splicing variants in various cancers represent crucial roles for tumorigenesis. Taken together, the identification of alternative spliced variants as biomarkers to distinguish between normal and cancer cells could cast light on tumorigenesis.

Keywords: alternative splicing, biological markers, neoplasms, neoplastic cell transformation

Introduction

Genes are transcribed into pre-mRNA and pass through splicing as a post-transcriptional modification to generate mature mRNA for translation [1]. Splicing is an essential process for gene expression in eukaryotes, eliminating introns and joining the exons, which occurs in the spliceosome, a complex of small nuclear ribonucleoproteins (snRNPs) and other proteins [2]. This splicing can generate alternative spliced transcripts from pre-mRNA and different exon constitutions, resulting in different proteins. Alternative splicing can occur in different ways. The most representative mechanisms are as follows: 1, extending or shortening the exon by alternative donor and acceptor sites; 2, exon skipping; 3, mutual exon exclusion; and 4, intron retention [1, 3, 4]. These mechanisms allow protein isoforms with different biological characters to be produced from single genes [1]. Therefore, complex transcriptomes and proteomes could be derived from a limited number of genes [4]. Namely, this process is an important strategy for the complicated regulation of eukaryotes, and most genes (92-95%) undergo this process [5, 6].

Alternative splicing is regulated according to cell type, developmental stage, and disease states [1, 7, 8]. Biochemical mechanisms for the recognition of splice sites are not understood clearly according to cellular conditions, but some tissue-specific factors participate in alternative splicing [4]. In addition, quantitative gene expression is controlled by nonsense-mediated decay mechanisms that degrade targeting mRNAs, producing nonsense mutations. Thus, truncated or erroneous proteins with abnormal functions are prevented from being expressed [9]. Alternative transcripts could be related to various diseases, including cancer. As many as 50% of genetic diseases of humans are related to mutations in splice site sequences and regulatory elements, such enhancers and silencers, resulting in alternative exon constitution [3, 10, 11]. Recently, the SpliceDisease database, providing information for relationships among gene mutation, splicing defects, and diseases, was reported [12]. Especially, aberrant spliced variants are found frequently in cancer, indicating that they could play a role for the survival of cancer cells [8]. Alternative splicing of cancer-related genes could affect cell cycle control, signal transduction pathways, apoptosis, angiogenesis, invasion,
Genetic and Epigenetic Regulation of Alternative Splicing

Alternative splicing is an essential post-transcriptional process for creating various protein isoforms from the same gene. Approximately 60% of human genes have at least one alternative splice transcript [15]. Alternative splicing is regulated by mutations in genetic regions [3, 7, 14]. Most exons are flanked by the intronic dinucleotides GT (donor site) and AG (acceptor site), which are recognized by the spliceosome (Fig. 1A). Pre-mRNA is reconstructed only by joined exons with the removal of introns. However, mutations of these splice sites lead to aberrant splicing, producing exon skipping and different joining of exons to truncated or nonfunctional or dysfunctional proteins. It was reported that mutations of splice sites may play important roles in human disease, and longer proteins also tend to associate to various disease [16, 17].

In the case of the PAX6 gene, a 5' splice site mutation in intron 12 induces exon skipping in relation to autosomal dominant aniridia [18]. In the APC gene, abnormal splicing by a 3' splice site mutation in intron 3 causes exon 4 skipping due to a frameshift in hepatoblastoma [19]. A mutation in intron 25 of the ABCA3 gene creates novel 5' splice sites [20]. The spliceosome is known to recognize cryptic splice sites instead of typical splice sites. Noncanonical splice sites are GC-AG, GG-AG, GT-TG, GT-CG, AT-AG, GA-AG, GT-AC, and CT-AG (5'-3' splice sites) [21-23]. Mutations at the noncanonical splice site of the SEDL gene cause variants in X-linked spondyloepiphysyeal dysplasia tarda [24]. Additionally, less exon skipping and more intron retention by

![Fig. 1. Genetic and epigenetic regulation of alternative splicing.](www.genominfo.org)
alternative splice sites are observed in cancer tissue compared to normal tissues from an analysis of expression sequence tags (ESTs) [25]. As a mutation of pre-mRNA, A to I RNA editing (conversion adenosine to inosine from deamination) also affects transcriptome diversification. RNA editing has regulatory roles, such as altering splice sites and sequences necessary for recognition of the spliceosome, resulting in modulation of alternatively spliced transcripts [26]. The last case of alternative splicing is by cis-genomic mutations in regulatory factors, such as branch sites, exonic and intronic splicing enhancers, and silencers [7, 27]. Pre-mRNA has exonic and intronic splicing enhancers (ESEs, ISEs) and silencers (ESSs, ISSs) that promote exon inclusion and exclusion by regulation of splice site recognition, respectively (Fig. 1A). The arginine-serine-rich (SR) protein that binds to ESEs induces splicing via a helping assembly spliceosome by interacting with snRNP. On the contrary, heterogeneous nuclear ribonucleoprotein is bound to ESEs and ISEs and inhibits splice site recognition by blocking spliceosome assembly [15, 27-29]. Mutations in regulatory factor could disturb the binding of these spliceosome assembly-related proteins. Mutations at nucleotide positions 57 and 58 of the 174-bp-long exon 7 cause exon 7 skipping as a result of aberrant splicing by interrupted ESE-specific consensus sequences that are recognized by the SC35 and SF2/ASF SR proteins [10, 30]. Additionally, Ron, encoding the tyrosine kinase receptor for macrophage-stimulating factor, is made from this process are developmentally regulated and aberrant alternative splicing occur frequently, leading to erroneous and dysfunctional proteins [14]. Protein isoforms made from this process are developmentally regulated and preferentially re-expressed in cancer and help the differentiation and survival of cancer cells (Fig. 2). The development}


domly in the genome. Exon skipping and mutually exclusive exons have significantly lower levels of both CG and mCG in the exonic regions, whereas intron retention has significantly higher levels of CG in both exonic and intronic regions [34]. A DNA-binding protein, CCCTC-binding factor (CTCF), was inhibited by a methylation event of CD45 exon 5 [39]. These epigenetic features are strongly associated with alternative splicing. Furthermore, these mechanisms are known to be changed according to cell type and disease states. Especially, in cancer, the epigenetic regulation of chromatin structure effects aberrant gene expression by alternative splicing in cancer [1, 7, 8].

Taken together, gene expression via alternative splicing is altered by complicated and mutual mechanisms, from genetic to epigenetic regulation. Therefore, it casts light on the understanding of cancer mechanisms by an investigation of alternative splicing patterns.

**Alternative Splicing in Cancer**

A number of studies have been reported that alternative splicing is related closely to development, cellular stress, and various diseases, including cancer, as a crucial contributor to transcriptome and proteome diversity [1, 7, 8]. In cancer, with increasing genomic instability, sequence substitution and aberrant alternative splicing occur frequently, leading to erroneous and dysfunctional proteins [14]. Protein isoforms made from this process are developmentally regulated and preferentially re-expressed in cancer and help the differentiation and survival of cancer cells (Fig. 2). The development
of genomewide analysis allows large-scale examination of the relationship between alternative splicing and tumorigenesis [40, 41]. Every case of alternative splicing is reported in cancer; among these, the most frequent case is the mutually exclusive exon [14]. For example, overexpression of CD44, involved in cell proliferation, differentiation, migration, and alternative splicing, by different splice sites of CD44 during tumorigenesis indicates that it could play roles in tumor cell invasion and metastasis [42-45]. Tumor suppressor genes, such as p53 and PTEN (Phosphatase and Tensin homolog, deleted on chromosome TEN), have splicing variants associated to cancer [46]. p53 protein isoforms through alternative splicing have critical roles in many biological processes, indicating that dysregulation effects tumorigenesis [47]. Different expression of PTEN and its alternatively spliced transcripts are found to vary in different tissue types. PTEN regulates p53 stability and in turn regulates its own transcriptional activity. The PTEN splice variants retained in intron 3 region and intron 5 region have been found in breast cancer [46]. In the case of the APC gene, aberrant splice skipping of exon 4, created by insertion of T in intron 4, is involved in colon cancer [48]. An alternative 5' splice site in BCL-X results in 2 isoforms, long and short (Bcl-x (L), Bcl-x (X)), which have contrasting functions related to apoptosis and are overexpressed in various tumors [49, 50]. Although specific roles of vascular endothelial growth factor (VEGF) isoforms are not known exactly, among the isoforms of VEGF4 involved in formation of new vessels, VEGF165 and VEGF165b, created by an alternative 3' splice site, have different expression in cancer [51, 52]. Over 40 different MDM2 transcripts are identified in normal and cancer by alternative 5' and 3' splice sites. They mostly lose the p53 binding domain and promote tumor progression and affect prognosis independently [53, 54]. Increasing number of reports have demonstrated the expression of aberrant and abnormal splice variants in cancer cells or tissues. However, it is not enough evidence for a functional relationship between alternative splicing and cancer. Hallmarks of the initiation and early growth of cancers during tumor progression using large-scale analyses of splicing variants in the cancer EST database are needed as new RNA prognosis markers.

**Table 1. Alternative spliced transcript related to cancer**

<table>
<thead>
<tr>
<th>Genes</th>
<th>Splice variants</th>
<th>Cancer types</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD44</td>
<td>Inclusion of exon</td>
<td>Various cancer</td>
<td>Tumor development and metastasis</td>
<td>Stickeler et al. [42]</td>
</tr>
<tr>
<td>p53</td>
<td>Retained intron</td>
<td>Breast cancer</td>
<td>Tumorigenesis</td>
<td>Okumura et al. [46]</td>
</tr>
<tr>
<td>p73</td>
<td>Inclusion of exon 13</td>
<td>Prostate cancer</td>
<td>Prostate hyperplasia and malignancy</td>
<td>Arvanitis et al. [55]</td>
</tr>
<tr>
<td>PTEN</td>
<td>Retained intron 3 and 5</td>
<td>Breast cancer</td>
<td>Pathogenesis of sporadic breast cancers with p53</td>
<td>Okumura et al. [46]</td>
</tr>
<tr>
<td>APC</td>
<td>Skipping of exon 4</td>
<td>Colon cancer</td>
<td>Cause of familial adenomatous polyposis (FAP)</td>
<td>Neklason et al. [48]</td>
</tr>
<tr>
<td>BCL-X</td>
<td>Alternative 5' splice site</td>
<td>Hepatocellular carcinoma</td>
<td>Regulation of apoptosis</td>
<td>Takehara et al. [50]</td>
</tr>
<tr>
<td>VEGF4</td>
<td>Alternative 3' splice site</td>
<td>Various cancer</td>
<td>Angiogenesis related to tumor progression</td>
<td>Pritchard-Jones et al. [52]</td>
</tr>
<tr>
<td>MDM2</td>
<td>Alternative 5' and 3' splice site</td>
<td>Various cancer</td>
<td>Promote tumor progression</td>
<td>Bartel et al. [54]</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Skipping of exon 18</td>
<td>Breast cancer</td>
<td>Breast cancer susceptibility</td>
<td>Okumura et al. [46]</td>
</tr>
<tr>
<td>TACC1</td>
<td>Inclusion of exon</td>
<td>Gastic cancer</td>
<td>Centrosome malfunction</td>
<td>Line et al. [56]</td>
</tr>
<tr>
<td>TERT</td>
<td>Alternative 3' splice site</td>
<td>Astrocytic gliomas</td>
<td>Loss of telomerase activity</td>
<td>Kotoula et al. [57]</td>
</tr>
<tr>
<td>KLF6</td>
<td>Alternative 5' splice site</td>
<td>Prostate cancer</td>
<td>Tumor cell proliferation</td>
<td>DiFeo et al. [58]</td>
</tr>
<tr>
<td>SURVIVIN</td>
<td>Retained intron</td>
<td>Various cancer</td>
<td>Tumorigenesis</td>
<td>Li [59]</td>
</tr>
<tr>
<td>ASP</td>
<td>Alternative 3' splice site</td>
<td>Hepatocellular carcinoma</td>
<td>Enhacement of Fas-mediated apoptosis</td>
<td>Hu et al. [60]</td>
</tr>
<tr>
<td>NF1</td>
<td>Alternative 5' and 3' splice site</td>
<td>Neurofibromatosis type I</td>
<td>Inactivation of tumor suppression</td>
<td>Barron and Lou [61]</td>
</tr>
<tr>
<td>Caspase 8</td>
<td>Retained intron</td>
<td>Breast cancer</td>
<td>Inhibitor of the caspase</td>
<td>Mohr et al. [62]</td>
</tr>
<tr>
<td>CDH17</td>
<td>Exclusion of exon 13</td>
<td>Hepatocellular carcinoma</td>
<td>Incidence of tumor recurrence</td>
<td>Wang et al. [63]</td>
</tr>
<tr>
<td>Ron</td>
<td>Skipping of exon 11</td>
<td>Colon cancer</td>
<td>Increasing of motility to expressing cells</td>
<td>Ghigna et al. [31]</td>
</tr>
</tbody>
</table>
tion (Fig. 2). For this reason, many studies have made an effort to identify specific features and regulation mechanisms of various diseases, including cancer. It is important and necessary to identify cancer markers to be able to distinguish between cancer and normal cells. Cancer markers could be very helpful in understanding tumorigenesis and developing tumor targets for therapeutic intervention. However, there still remain unsolved problems in spite of a number of studies [7, 14, 45, 64]. From this point of view, alternatively spliced transcripts by mutation and cryptic splice site altered expression levels in cancer have emerged as strong candidates of cancer biomarker at the mRNA and protein level [7]. As written previously, some alternative transcripts of several genes, CD44, p53, PTEN, BCL-X, VEGF4, and MDM2, have been discovered (see also Table 1) [43, 46, 49, 50, 52, 54]. They are also associated with various cancers directly or indirectly. Recently, the discovery of biomarkers has improved by genomewide analysis [13, 65, 66]. It could be expected to provide valuable information for the association between alternative splice variants and cancers. Thus, novel candidate variants could contribute to the development of diagnostic, prognostic, and therapeutic markers.

**Conclusion**

RNA splicing is a core mechanism to generate mature mRNA for translation, and alternative splicing is an indispensable mechanism, strategically leading to protein diversity for complicated regulation in eukaryotes. In cancer, alternative splicing is more flexible, leading various proteins with aberrant functions to promote growth and the spread of cancer cells. It is important to identify alternative transcripts that function specifically in cancer. These alternative transcripts could be used not only as diagnostic biomarkers but also prognostic and therapeutic biomarkers. Therefore, studies on genetic and epigenetic regulation in relation to alternative splice variants in cancer could open new windows of research in answering unsolved questions of tumorigenesis.

**Acknowledgments**

This work was supported by a 2-Year Research Grant of Pusan National University.

**References**


33. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of alternative splicing by histone modifications. *Science* 2010;327:996-1000.


58. DiFeo A, Martignetti JA, Narla G. The role of KLF6 and its


