



Chapter B07. Alternative Splicing: Lessons from Cancer

Sandro J. de Souza¹ and Anamaria A. Camargo²

Created: October 12, 2006; Updated: September 17, 2007.

Introduction

The last 15 years have witnessed an impressive development of a discipline that uses computational expertise and resources to approach biological problems. Bioinformatics, or computational biology as we know this discipline nowadays, is one of the pillars of biomedical sciences in the 21st century. Similar to what happened to molecular biology in the 1980s, it is expected that bioinformatics/computational biology will become a widespread discipline. Basically, all research groups will need to have their own computational expertise to explore the vast amount of data available in the public domain and to integrate that extracted knowledge into their own research program.

How did this happen? It was simply a consequence of molecular biology. First, the molecular biology revolution allowed the generation of huge amounts of data, illustrated by all the genome projects finished in the last 10 years. From the first moment, biologists needed computers to handle the data. More recently, biologists have used computational tools to make “experiments” and infer biological and functional significance from knowledge derived from *in silico* analyses. Comparative genomics is an example of such rationale.

One of the fields that has progressed more effectively under this new paradigm is transcriptomics. In the early 1990s, the first large-scale approach was used to generate expression data in the form of expressed sequence tags (ESTs) (1). As of June 2007, dbEST contains more than 43 million ESTs from a variety of organisms (dbEST release 062207: <http://www.ncbi.nlm.nih.gov/dbEST>). More recently, SAGE (2), MPSS (3), and microarray (4) technologies have added a more quantitative character to the collection of expression data. The completion of the Human Genome Project provided a scaffold onto which the expression data were mapped. The human genome sequence, now “decorated” with many different types of data, has become a driving resource in the biomedical sciences.

In this review, we will discuss how bioinformatics/computational biology and the genome sequence have been used by us and others to study the phenomenon of alternative splicing. We will give special emphasis to the association between alternative splicing and cancer.

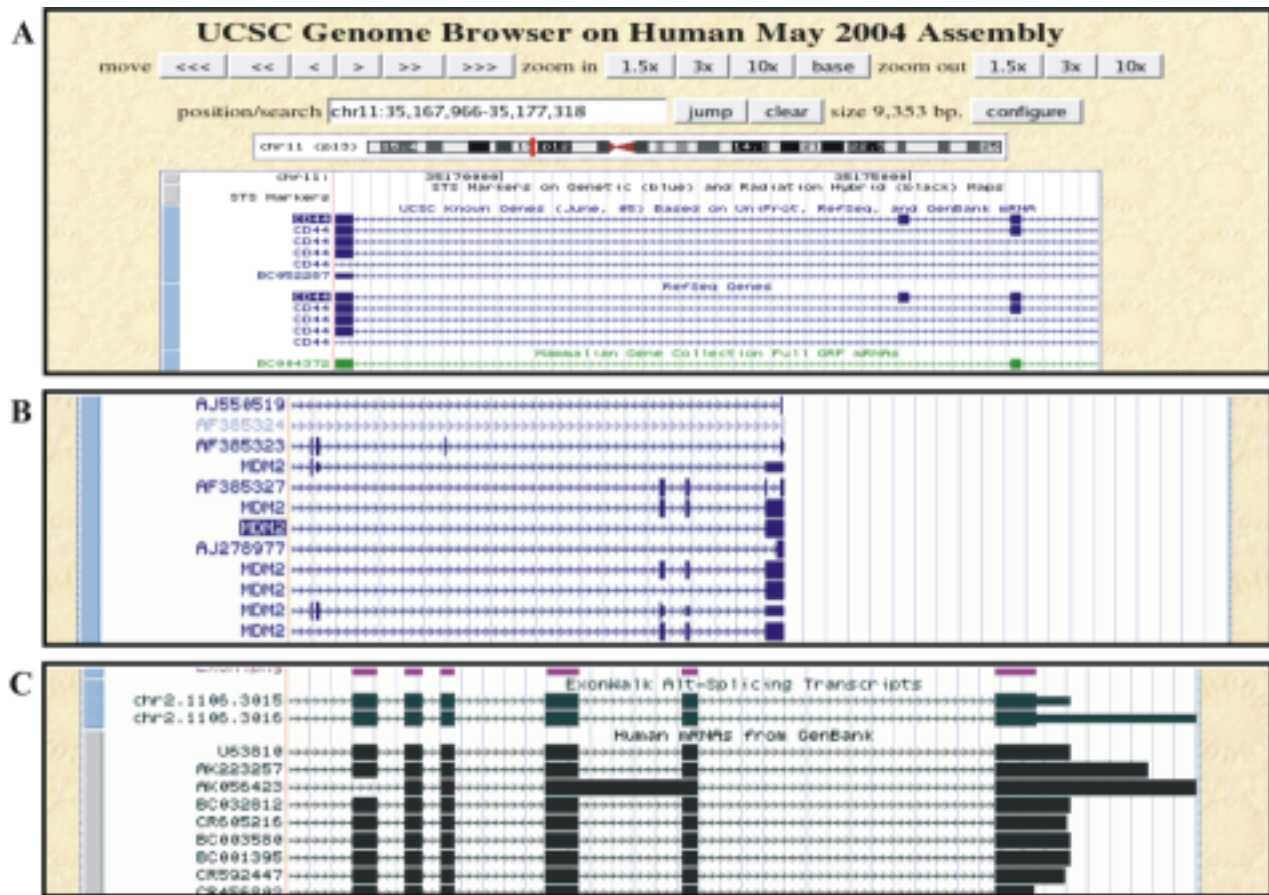


Figure 1. UCSC Genome Browser showing the three major types of alternative splicing. A, exon skipping reported for gene *CD44*. B, usage of alternative acceptor site for gene *MDM2*. One of the alternative acceptor sites is reported by sequence AJ278977 (last exon). C, intron retention reported for gene *WDR39*. The retention event is reported by sequence AK056423.

Alternative Splicing

The splicing of introns in eukaryotic genes is one of the most basic processes within the cell. We are still, however, far from a complete understanding of the regulatory mechanisms acting on splicing. From a chemical standpoint, splicing is very simple. It corresponds to two *trans*-esterification reactions that remove an intervening sequence (the intron) and join the two flanking exons. Biologically, however, splicing is a complex and intricate process. A huge ribonucleoprotein complex, the spliceosome (containing five RNAs and hundreds of proteins), is needed. In spite of the fact that introns are on average 10 times longer than exons, the spliceosome recognizes the correct intron/exon borders with an astonishing precision. This is achieved by a balanced mix of *cis* and *trans* elements. To make things even more complex, elements in *cis* are short and weak. Some of these sequence elements are present in basically all introns: a donor site (|GT), an acceptor site (AG|), a branch site, and a polypyrimidine tract (see Ast (5) and Woodley and Valcarcel (6) for reviews). These elements *per se* are not informative enough to make sure that cells can recognize the correct borders with precision. There are additional sequence elements, most unknown, that are probably binding sites for RNAs and/or proteins—the *trans* factors. These elements in *trans* are in the spliceosome and bind the elements in *cis* positioning in the radical groups involved in the splicing reaction. Some of these elements can enhance splicing, such as the exonic or intronic splicing enhancers (ESEs or ISEs, respectively). On the other hand, they can also silence splicing and, therefore, are called exonic or intronic splicing silencers (ESSs or ISSs, respectively). Several groups of proteins are now known to bind these elements. SR proteins, for example, bind enhancers preferentially, whereas hnRNPs bind silencers preferentially (for a review, see Matlin *et al.* (7)).

As predicted by Wally Gilbert (8) in his seminal “Why genes in pieces?” paper in the late 1970s, alternative exon/intron borders can be used by cells with profound implications for the encoded protein. In fact, splicing variants have been found since then for a variety of genes in several species. As shown in Figure 1, there are several types of alternative splicing. The most common involves the skipping of one or more exons. Alternative donor and acceptor sites can also be used, and finally, introns can be retained in the mature message.

The availability of a large amount of expressed data in the form of ESTs has allowed large-scale studies on alternative splicing in mammals, especially mouse and human (9, 10). Surprisingly, these analyses showed that alternative splicing is much more frequent than originally estimated. At least one-half of all human genes undergo alternative splicing, and this number is certainly underestimated, because for those genes with a high expression level and consequently more represented in dbEST, the rate is close to 90%. This high frequency of alternative splicing has raised concerns about the biological significance of the splicing variants. Although there is evidence for a functional significance of some splicing variants, some authors have argued that a significant fraction of all splicing variants is spurious. These variants could be simply the products of leaking in the splicing reaction. In fact, a significant amount of splicing variants occurring in the coding region of genes does not maintain the respective reading frame (11). However, if we take those variants that are present in both mouse and human, most of them conserve the reading frame. Interestingly, some have argued that these spurious variants can have a functional role by down-regulating the normal function of a gene. These authors have termed this process RUST (regulated unproductive splicing and translation) (12, 13). These “unproductive” messages would be degraded by a process known as nonsense-mediated decay (NMD) (14). NMD is an RNA surveillance system that recognizes and targets for destruction those messages presenting a premature stop codon.

Also important in this field is the characterization of the regulation of alternative splicing. Much of our knowledge about the intricate mechanisms regulating constitutive and alternative splicing was achieved through individual laboratories working on specific models. More recently, large-scale approaches have been used extensively in the identification of putative regulatory elements. For example, proteomic analysis has been used in the identification of proteic components of the spliceosome (15). SELEX has been used extensively in the characterization of the RNA-binding specificity of many splicing factors (16). Finally, a combination of experimental and bioinformatics approaches has been used extensively by Chris Burge’s group in the identification of different types of regulatory elements (17-19).

Computational Approaches to Detect Splicing Variants

Almost the totality of data on splicing variants was deduced by different methods involving sequence comparisons. The idea behind these methodologies is that splicing variants from the same gene share some common sequence and can therefore be grouped together. ESTs are the major source in these studies aimed to catalog the splicing variants. Mironov *et al.* (20), for example, used EST contigs from the TIGR gene index to make an inventory of intron–exon structures in the set of known human genes. With the availability of the human genome sequence, several groups started to make more precise inferences about alternative splicing by simply mapping all cDNAs onto the genome sequence (9, 21). The use of more precise algorithms of alignment, such as Sim4 (22), which pays special attention to the exon/intron border, made these inferences more reliable.

Instead of comparing pairwise alignments, the state-of-art methodology nowadays aligns all cDNAs against the genome, loads the cDNA coordinates into relational databases, and compares the borders among different cDNAs from the same gene. Details on the approach used by our group can be found elsewhere (21, 23, 24). Figure 2 shows two variants from gene *WDR39* extracted directly from our relational database.

The major advantage of this approach is the possibility to perform genome-wide analyses through the development of software that interrogates the relational database. The knowledge extracted can be inserted back

```
mysql> select * from cdna_blocks where seqacc="AK056423";
```

contigid	seqacc	indexid	blocknum	blockid	start_cdna	end_cdna	start_gnm	end_gnm	strand	ident	cluster
CHR2	AK056423	1	0	NULL	1	250	96353852	96354101	+	99	Hs.12109
CHR2	AK056423	1	0	NULL	251	362	96355237	96355348	+	99	Hs.12109
CHR2	AK056423	1	0	NULL	363	451	96355448	96355536	+	100	Hs.12109
CHR2	AK056423	1	0	NULL	452	1350	96356069	96356967	+	99	Hs.12109
CHR2	AK056423	1	0	NULL	1351	3902	96358723	96361274	+	99	Hs.12109

```
5 rows in set (0.00 sec)
```

```
mysql> select * from cdna_blocks where seqacc="BC032812";
```

contigid	seqacc	indexid	blocknum	blockid	start_cdna	end_cdna	start_gnm	end_gnm	strand	ident	cluster
CHR2	BC032812	1	0	NULL	1	238	96353864	96354101	+	100	Hs.12109
CHR2	BC032812	1	0	NULL	239	387	96354933	96355081	+	100	Hs.12109
CHR2	BC032812	1	0	NULL	388	499	96355237	96355348	+	100	Hs.12109
CHR2	BC032812	1	0	NULL	500	588	96355448	96355536	+	100	Hs.12109
CHR2	BC032812	1	0	NULL	589	790	96356069	96356270	+	100	Hs.12109
CHR2	BC032812	1	0	NULL	791	878	96356880	96356967	+	100	Hs.12109
CHR2	BC032812	1	0	NULL	879	1326	96358723	96358170	+	100	Hs.12109

```
7 rows in set (0.00 sec)
```

```
mysql>
```

Figure 2. Structure of an in-house database reporting two splicing variants for gene *WDR39*. The first query for sequence AK056423 reports five exons. The second query for sequence BC032812 reports seven exons. Differences are attributable to an inclusion of the second exon in the sequence BC032812 and an intron retention in sequence AK056423 (fourth exon in AK056423 corresponds to the fifth and sixth exons in BC032812).

in the same database in the form of different tables. The information can also be easily linked to other types of data because most of the related information currently uses the genome sequence as a scaffold.

Differential Expression of Splicing Variants in Tumors

Alternative splicing has been involved in many different biological phenomena including sex determination in *Drosophila* (25), among others. As expected, alternative splicing seems to be quite important in the pathogenesis of several human diseases. It is believed that around 15% of all human genetic diseases are caused by mutations in sequence elements important for constitutive splicing. One of the first splicing mutations described activates a cryptic acceptor site in the β -globin gene resulting in β^+ -thalassemia. The involvement of splicing variants in a specific disease can be more subtle. Let's take, as an example, the fronto-temporal dementia with Parkinson (FTDP-17), a neurological disease linked to chromosome 17 (26). One of the first candidate genes for this disease was the protein tau, a microtubule-associated protein involved with axonal transport in neurons. Indeed, mutations in tau have been reported to be associated with development of FTDP-17 (27). More recently, mutations that alter the ratio of splicing variants, skipping or including exon 10, have also been associated with FTDP-17 (28).

Isolated cases of differential expression of splicing variants in cancer have been reported in the last 10 years (for a review, see Caballero *et al.* (29)). For example, Bcl-x, an apoptosis regulator, has two splicing variants because of alternative donor sites in its exon 2. Only the longer form is differentially expressed in small cell lung carcinoma (30) and breast carcinoma (31). The most known example, however, is CD44, a cell surface glycoprotein. Differential expression of several splicing variants has been observed for a range of different tumors (for a review, see Caballero *et al.* (29)). More recently, Matsushita *et al.* (32) reported that a splicing variant of FIR (FUSE-binding protein-interacting repressor) is unable to repress c-Myc and to drive apoptosis. This splicing variant was only expressed in colorectal cancer cells and was not detected in the adjacent normal cells. The results presented in this report suggest that this variant promotes tumor development. Narla *et al.* (33)

reported an association between a germline polymorphism with both an unbalanced expression of a splicing variant of KLF6 and an increased risk for prostate cancer.

The increasing amount of cDNA libraries constructed from both normal and tumor samples allows the development of genome-wide strategies for the identification of tumor-associated splicing variants. Several groups reported genome-wide screening strategies searching for splicing variants differentially expressed in tumors (23, 34-36). Without exception, these authors made use of the huge amount of cDNA data available in the public databases to identify putative variants differentially expressed in tumors. The proportional frequency of cDNAs derived from distinct variants is an indication whether a given variant is differentially expressed in a library or in a pool of libraries. One of the major problems affecting this type of analysis is the identification of genes, not variants, differentially expressed in tumors. This happens because the computational and statistical methods used in the analyses are not sensitive enough to discriminate the expression level of all variants from a given gene. More recently, we tried to overcome this limitation by using SAGE data to discriminate variants differentially expressed from genes differentially expressed (23). Our computational approach identified more than 1,300 splicing variants putatively associated with cancer. Experimental validation for a subset of these candidates was achieved for both tumor cell lines and patient samples.

All of these reports reinforce the notion that cancer cells reprogram the splicing pattern of their genes. How? One possibility is that splicing factors are differentially expressed in tumors, and this causes a downstream effect on the transcriptome of these cells. We have recently evaluated this possibility through the use of SAGE and microarray data (37). We showed that splicing factors are indeed differentially expressed in tumors.

The discovery of splicing variants associated with cancer represents a promising strategy in the fight against this terrible disease. As already discussed, changes in splicing have been shown to play a functionally significant role in tumorigenesis. They can also serve as targets for the early diagnosis of cancer. Furthermore, cancer-specific splicing variants can present new epitopes recognized by the immune system and may serve as targets for immunotherapy.

Large-Scale Analysis of Alternative Splicing in Parasites

Parasites are a diverse group of organisms that share the common features of living within and exploiting a host organism. Most parasites undergo complex, multiphase life cycles that can involve extracellular and intracellular stages as well as different hosts. As a result of this complexity and unusual approaches to survival, parasites developed some remarkable adaptations at both the genetic and biochemical level that allow them, for example, to evade host defenses and to facilitate transmission to new hosts.

Molecular studies of parasites, particularly of trypanosomatids and apicomplexans, have revealed novel aspects of gene regulation and expression that could later be applicable to higher eukaryotes. Post-transcriptional modifications, such as *trans*-splicing (38) and RNA editing (39), were first observed and characterized at the molecular level in trypanosomatids. *cis*-Splicing also occurs among parasites, although to a lesser extent than that observed for higher eukaryotes. Interesting examples of alternative splicing in parasites have already been documented. Alternative splicing isoforms of the hypoxanthine-xanthine-guanine phosphoribosyltransferase gene (*HXGPRT*) have been observed in *Toxoplasma gondii*. These isoforms differ in the presence or absence of a 49-amino acid insertion (which is specified by a single, differentially spliced exon) and present different cellular localization, suggesting the existence of functional differences between both isoforms (40). Alternative splicing isoforms of the *PfPK6* gene in *Plasmodium falciparum* have also been observed and shown to be differentially expressed during different asexual erythrocytic stages of this parasite (41).

These initial works suggest that alternative splicing might have an important role in parasites' biology and could regulate important events such as invasion into host cells and evasion of host defense. Genome sequencing projects have been carried or are under way for several relevant human parasites, such as *P. falciparum* and *P.*

vivax, *Trichomonas vaginalis*, *Toxoplasma gondii*, *Schistosoma mansoni*, *Leishmania major*, *Trypanosoma cruzi*, and *T. brucei* (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome>). Expressed sequences (ESTs and full-length cDNA clones) from different life stages of most of these parasites are also publicly available (<http://www.tigr.org/tdb/tgi/index.shtml>). In this context, large-scale analysis of alternative splicing in parasites, using computational tools similar to the ones we described in this chapter, would be extremely important for a better understanding of parasites' biology and could eventually have important practical implications for preventing, diagnosing, and treating parasitic diseases.

Final Comments

As discussed by Wang *et al.* (19), the characterization of an “RNA splicing code” will require a detailed catalog of all splicing variants, all splicing factors, all regulatory elements, and the interactions among all of these elements. We expect that bioinformatics/computational biology will continue to provide essential information for the completion of this catalog and finally, a full characterization of alternative splicing.

References

1. Adams MD, Dubnick M, Kerlavage AR, Moreno R, Kelley JM, Utterback TR, Nagle JW, Fields C, Venter JC. Sequence identification of 2,375 human brain genes. *Nature*. 1992;355:632–634. PubMed PMID: 1538749.
2. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science*. 1995;270:484–487. PubMed PMID: 7570003.
3. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao J, Corcoran K. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol*. 2000;18:630–634. PubMed PMID: 10835600.
4. Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D. Light-directed, spatially addressable parallel chemical synthesis. *Science*. 1991;251:767–773. PubMed PMID: 1990438.
5. Ast G. How did alternative splicing evolve? *Nat Rev Genet*. 2004;5:773–782. PubMed PMID: 15510168.
6. Woodley L, Valcarcel J. Regulation of alternative pre-mRNA splicing. *Brief Funct Genomic Proteomic*. 2002;1:266–277. PubMed PMID: 15239893.
7. Matlin AJ, Clark F, Smith CW. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol*. 2005;6:386–398. PubMed PMID: 15956978.
8. Gilbert W. Why genes in pieces? *Nature*. 1978;271:501. PubMed PMID: 622185.
9. Modrek B, Lee CJ. Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat Genet*. 2003;34:177–180. PubMed PMID: 12730695.
10. Xu Q, Modrek B, Lee C. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res*. 2002;30:3754–3766. PubMed PMID: 12202761.
11. Resch A, Xing Y, Alekseyenko A, Modrek B, Lee C. Evidence for a subpopulation of conserved alternative splicing events under selection pressure for protein reading frame preservation. *Nucleic Acids Res*. 2004;32:1261–1269. PubMed PMID: 14982953.
12. Hillman RT, Green RE, Brenner SE. An unappreciated role for RNA surveillance. *Genome Biol*. 2004;5:R8. PubMed PMID: 14759258.
13. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A*. 2003;100:189–192. PubMed PMID: 12502788.
14. Morrison M, Harris KS, Roth MB. smg mutants affect the expression of alternatively spliced SR protein mRNAs in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 1997;94:9782–9785. PubMed PMID: 9275202.
15. Rappsilber J, Ryder U, Lamond AI, Mann M. Large-scale proteomic analysis of the human spliceosome. *Genome Res*. 2002;12:1231–1245. PubMed PMID: 12176931.
16. Tacke R, Manley JL. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J*. 1995;14:3540–3551. PubMed PMID: 7543047.

17. Fairbrother WG, Yeo GW, Yeh R, Goldstein P, Mawson M, Sharp PA, Burge CB. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* 2004;32:W187-90. PubMed PMID: 15215377.
18. Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science.* 2002;297:1007–1013. PubMed PMID: 12114529.
19. Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB. Systematic identification and analysis of exonic splicing silencers. *Cell.* 2004;119:831–845. PubMed PMID: 15607979.
20. Mironov AA, Fickett JW, Gelfand MS. Frequent alternative splicing of human genes. *Genome Res.* 1999;9:1288–1293. PubMed PMID: 10613851.
21. Galante PA, Sakabe NJ, Kirschbaum-Slager N, de Souza SJ. Detection and evaluation of intron retention events in the human transcriptome. *RNA.* 2004;10:757–765. PubMed PMID: 15100430.
22. Florea L, Hartzell G, Zhang Z, Rubin GM, Miller W. A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.* 1998;8:967–974. PubMed PMID: 9750195.
23. Kirschbaum-Slager N, Parmigiani RB, Camargo AA, de Souza SJ. Identification of human exons overexpressed in tumors through the use of genome and expressed sequence data. *Physiol Genomics.* 2005;21:423–432. PubMed PMID: 15784694.
24. Sakabe NJ, de Souza JE, Galante PA, de Oliveira PS, Passetti F, Brentani H, Osorio EC, Zaiats AC, Leerkes MR, Kitajima JP, Brentani RR, Strausberg RL, Simpson AJ, de Souza SJ. ORESTES are enriched in rare exon usage variants affecting the encoded proteins. *C R Biol.* 2003;326:979–985. PubMed PMID: 14744104.
25. Baker BS. Sex in flies: the splice of life. *Nature.* 1989;340:521–524. PubMed PMID: 2505080.
26. Lynch T, Sano M, Marder KS, Bell KL, Foster NL, Defendini RF, Sima AA, Keohane C, Nygaard TG, Fahn S, et al. Clinical characteristics of a family with chromosome 17-linked disinhibition-dementia-parkinsonism-amyotrophy complex. *Neurology.* 1994;44:1878–1884. PubMed PMID: 7936241.
27. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, Pickering-Brown S, Chakraverty S, Isaacs A, Grover A, Hackett J, Adamson J, Lincoln S, Dickson D, Davies P, Petersen RC, Stevens M, de Graaff E, Wauters E, van Baren J, Hillebrand M, Joosse M, Kwon JM, Nowotny P, Che LK, Norton J, Morris JC, Reed LA, Trojanowski J, Basun H, Lannfelt L, Neystat M, Fahn S, Dark F, Tannenberg T, Dodd PR, Hayward N, Kwok JB, Schofield PR, Andreadis A, Snowden J, Craufurd D, Neary D, Owen F, Oostra BA, Hardy J, Goate A, van Swieten J, Mann D, Lynch T, Heutink P. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature.* 1998;393:702–705. PubMed PMID: 9641683.
28. Grover A, DeTure M, Yen SH, Hutton M. Effects on splicing and protein function of three mutations in codon N296 of tau in vitro. *Neurosci Lett.* 2002;323:33–36. PubMed PMID: 11911984.
29. Caballero OL, de Souza SJ, Brentani RR, Simpson AJ. Alternative spliced transcripts as cancer markers. *Dis Markers.* 2001;17:67–75. PubMed PMID: 11673653.
30. Reeve JG, Xiong J, Morgan J, Bleehen NM. Expression of apoptosis-regulatory genes in lung tumor cell lines: relationship to p53 expression and relevance to acquired drug resistance. *Br J Cancer.* 1996;73:1193–1200. PubMed PMID: 8630278.
31. Olopade OI, Adeyanju MO, Safa AR, Hagos F, Mick R, Thompson CB, Recant WM. Overexpression of BCL-x protein in primary breast cancer is associated with high tumor grade and nodal metastases. *Cancer J Sci Am.* 1997;3:230–237. PubMed PMID: 9263629.
32. Matsushita K, Tomonaga T, Shimada H, Shioya A, Higashi M, Matsubara H, Harigaya K, Nomura F, Libutti D, Levens D, Ochiai T. An essential role of alternative splicing of *c-myc* suppressor FUSE-binding protein-interacting repressor in carcinogenesis. *Cancer Res.* 2006;66:1409–1417. PubMed PMID: 16452196.
33. Narla G, DiFeo A, Reeves HL, Schaid DJ, et al. A germline DNA polymorphism enhances alternative splicing of the KLF6 tumor suppressor gene and is associated with increased prostate cancer risk. *Cancer Res.* 2005;65:1213–1222. PubMed PMID: 15735005.
34. Hui L, Zhang X, Wu X, Lin Z, Wang Q, Li Y, Hu G. Identification of alternatively spliced mRNA variants related to cancers by genome-wide ESTs alignment. *Oncogene.* 2004;23:3013–3023. PubMed PMID: 15048092.

35. Wang Z, Lo HS, Yang H, Gere S, Hu Y, Buetow KH, Lee MP. Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer. *Cancer Res.* 2003;63:655–657. PubMed PMID: 12566310.
36. Xu Q, Lee C. Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences. *Nucleic Acids Res.* 2003;31:5635–5643. PubMed PMID: 14500827.
37. Kirschbaum-Slager N, Lopes GM, Galante PA, Riggins GJ, de Souza SJ. Splicing factors are differentially expressed in tumors. *Genet Mol Res.* 2004;3:532–544. PubMed PMID: 15688319.
38. Liang XH, Haritan A, Uliel S, Michaeli S. trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot Cell.* 2003;2(5):830–840. PubMed PMID: 14555465.
39. Stuart K, Panigrahi AK. RNA editing: complexity and complications. *Mol Microbiol.* 2002;45(3):591–596. PubMed PMID: 12139607.
40. Chaudhary K, Donald RG, Nishi M, Carter D, Ullman B, Roos DS. Differential localization of alternatively spliced hypoxanthine-xanthine-guanine phosphoribosyltransferase isoforms in *Toxoplasma gondii*. *J Biol Chem.* 2005;280(23):22053–22059. PubMed PMID: 15814612.
41. Bracchi-Ricard V, Barik S, Delvecchio C, Doerig C, Chakrabarti R, Chakrabarti D. PfPK6, a novel cyclin-dependent kinase/mitogen-activated protein kinase-related protein kinase from *Plasmodium falciparum*. *Biochem J.* 2000;347(Pt 1):255–263. PubMed PMID: 10727426.