

Alternative spliced transcripts as cancer markers

Otavia L. Caballero^b, Sandro J. de Souza^b,
Ricardo R. Brentani^{a,b} and
Andrew J.G. Simpson^{b,*}

^a*Hospital do Cancer AC Camargo, São Paulo, Brazil*

^a*Ludwig Institute for Cancer Research, São Paulo, Brazil*

Eukaryotic mRNAs are transcribed as precursors containing their intronic sequences. These are subsequently excised and the exons are spliced together to form mature mRNAs. This process can lead to transcript diversification through the phenomenon of alternative splicing. Alternative splicing can take the form of one or more skipped exons, variable position of intron splicing or intron retention. The effect of alternative splicing in expanding protein repertoire might partially underlie the apparent discrepancy between gene number and the complexity of higher eukaryotes. It is likely that more than 50% of human genes produce more than one transcript form. Many cancer-associated genes, such as CD44 and WT1 are alternatively spliced. Variation of the splicing process occurs during tumor progression and may play a major role in tumorigenesis. Furthermore, alternatively spliced transcripts may be extremely useful as cancer markers, since it appears likely that there may be striking contrasts in usage of alternatively spliced transcript variants between normal and tumor tissue than in alterations in the general levels of gene expression.

1. Introduction

The improved management of human cancer will depend on early detection, more accurate prognostic assessment and the availability of a larger selection of therapeutic agents. There is no doubt that the current intense exploitation of the human transcriptome will make a fundamental contribution in each of these areas. We know that cancer is the result of surprisingly subtle

changes in overall gene expression caused both by mutation of key regulatory genes or epigenetic phenomena such as aberrant methylation. Quantification of gene expression using microarrays or the serial analysis of gene expression (SAGE) estimate that only some 5% of genes have altered expression levels in tumors as compared with corresponding normal tissues. The percentages that exhibit down regulation and up regulation are approximately equal. Moreover, the majority of differentially expressed genes encodes proteins associated with basic cellular functions and in particular reflects the altered proliferative state of the malignant cell. They are thus not causal agents of the tumor nor attractive drug targets. Alteration of the levels of expression is also on the whole rather modest with very few genes consistently showing more than 5-fold differences between normal and tumor samples. These limited alterations render the use of these genes as tumor markers requiring sensitive and accurate detection techniques as well as very well defined tissue samples. Many human genes remain to be identified due to the still draft stage of the available human genome sequence, the impossibility of accurate gene prediction using currently available computational tools and the lack of sufficient transcript sequence data. Microarray analysis, in particular, is only as useful as their DNA content permits. Thus as we complete the inventory of human genes and locate them in microarrays, examples more fundamentally transcriptionally regulated may be discovered. These are likely to be rather poorly expressed genes, however, as most genes that exhibit higher levels of expression are already fully defined. Again these changes will be relatively difficult to use in clinical assays.

The complete definition of the human transcriptome will not only permit the eventual global assessment of differential gene expression but will also lead to the identification of all the different transcript forms that can originate from the same gene. It is possibly in this area that the most exciting prospects for harnessing the power of genomics in the fight against cancer lie. Possibly the majority of genes in the human genome pro-

* Address for correspondence: Dr. Andrew J.G. Simpson, Ludwig Institute for Cancer Research, Rua Prof. Antônio Prudente 109 – 4o andar, São Paulo, SP 01509-010, Brazil. Tel.: +55 11 270 4922; Fax: +55 11 270 7001; E-mail: asimpson@node1.com.br.

duce alternative transcripts that when transcribed lead to protein products with altered structure and function. Crucially there is already evidence that such alternative transcript generation, through the process of alternative splicing as described below, can accompany the process of tumorigenesis. Indeed, the generation of alternatively spliced transcripts may prove to be one of the causal steps in the development of some cancers. In this instance, there may be complete presence or absence of particular splicing forms in a tumor as opposed to a corresponding normal tissue rendering them highly suitable both as cancer markers and potential drug targets. Although this area of research still remains to be explored in any detail, indeed it is likely that the vast majority of human transcripts types still remain unknown to us, the preliminary data available raise exciting possibilities.

2. Alternative transcript splicing

Eukaryotic mRNAs are transcribed as precursors containing their intronic sequences. Subsequently, these are excised and the exons are spliced together to form mature mRNAs. This process can lead to diversification through the phenomenon of alternative splicing. Variation in mRNA structure takes many different forms [1,2] that include the use of cryptic donor and acceptor splicing sites, exon skipping and the use of intronic sequence as an exon. Also, the positions of either 5' or 3' splice sites can shift to make exons longer or shorter. In addition to these changes in splicing, alterations in the transcriptional start site or the polyadenylation site also allow production of multiple mRNAs from a single gene (Fig. 1). The consequences of alternative splicing range from switching expression of a protein on and off, by excluding and including stop codons, to structural and functional diversification of protein products. Usually, the process is highly regulated so that particular splicing patterns occur only under particular conditions. It is becoming clear that alternative splicing has an extremely important role in expanding the protein repertoire and might therefore partially underlie the apparent discrepancy between gene number and the complexity of higher eukaryotes. Indeed, alternative splicing can generate more transcripts from a single gene than the total number of genes in an entire genome [3]. As yet, however, for the vast majority of alternative splicing events their functional significance remains unknown [4].

The effect of altered mRNA splicing on the structure of the encoded protein can be profound [1,2]. In some transcripts, whole functional domains are added to or subtracted from the protein coding sequence. In other systems, the introduction of an early stop codon can result in a truncated protein, transforming a membrane bound into a soluble protein, for example, or an unstable mRNA. Alternative splicing is also commonly used to control the inclusion of particular short peptides sequences within a longer protein. These optional sequence cassettes range from one to hundreds of aminoacids in length, and have many specific effects on the activity of a protein product. Changes in splicing have been shown to determine the ligand binding of growth factor receptors and cell adhesion molecules, and to alter the activation domains of transcriptional factors [1,2]. Furthermore, the splicing pattern of an mRNA may determine the subcellular localization of the encoded protein, the phosphorylation of proteins by kinases or the binding of an enzyme by its allosteric effector. Determining how these sometimes-subtle changes in the sequence affect protein function is a crucial question in many different problems in developmental and cell biology, including control of apoptosis, neuronal connectivity, cell contraction and tumor progression [4].

The amount of variation that can be generated from a single gene through alternative splicing is astonishing. An example is the *Drosophila melanogaster* DSCAM gene [5]. Each transcript of the gene contains 24 exons that encode an axon guidance receptor. However, the gene contains an array of potential alternatives for exons 4,6,9 and 17. These exons are used in a mutually exclusive way with 12 alternatives for exon 4, 48 for exon 6, 33 for exon 9 and 2 for exon 17. Thus, alternative splicing can potentially generate more than 38,000 transcripts from this gene!

How do cells choose between specific splicing pathways? The information content in a splicing site is limited. The most accurate computer programs achieve an identification rate of approximately 50% [2]. Mutations that destroy splice sites or create new ones are responsible for 15% of all human genetic disease. The initial commitment of an mRNA to splicing involves a series of interactions between the mRNA and several other RNAs and proteins. A family of proteins critical to splicing is the serine-arginine family of splicing factors (SR proteins) [6]. These proteins seem to act as bridges between the mRNA and several other protein factors.

Alignment of both full length transcripts and expressed sequence tags (ESTs) has provided a minimum

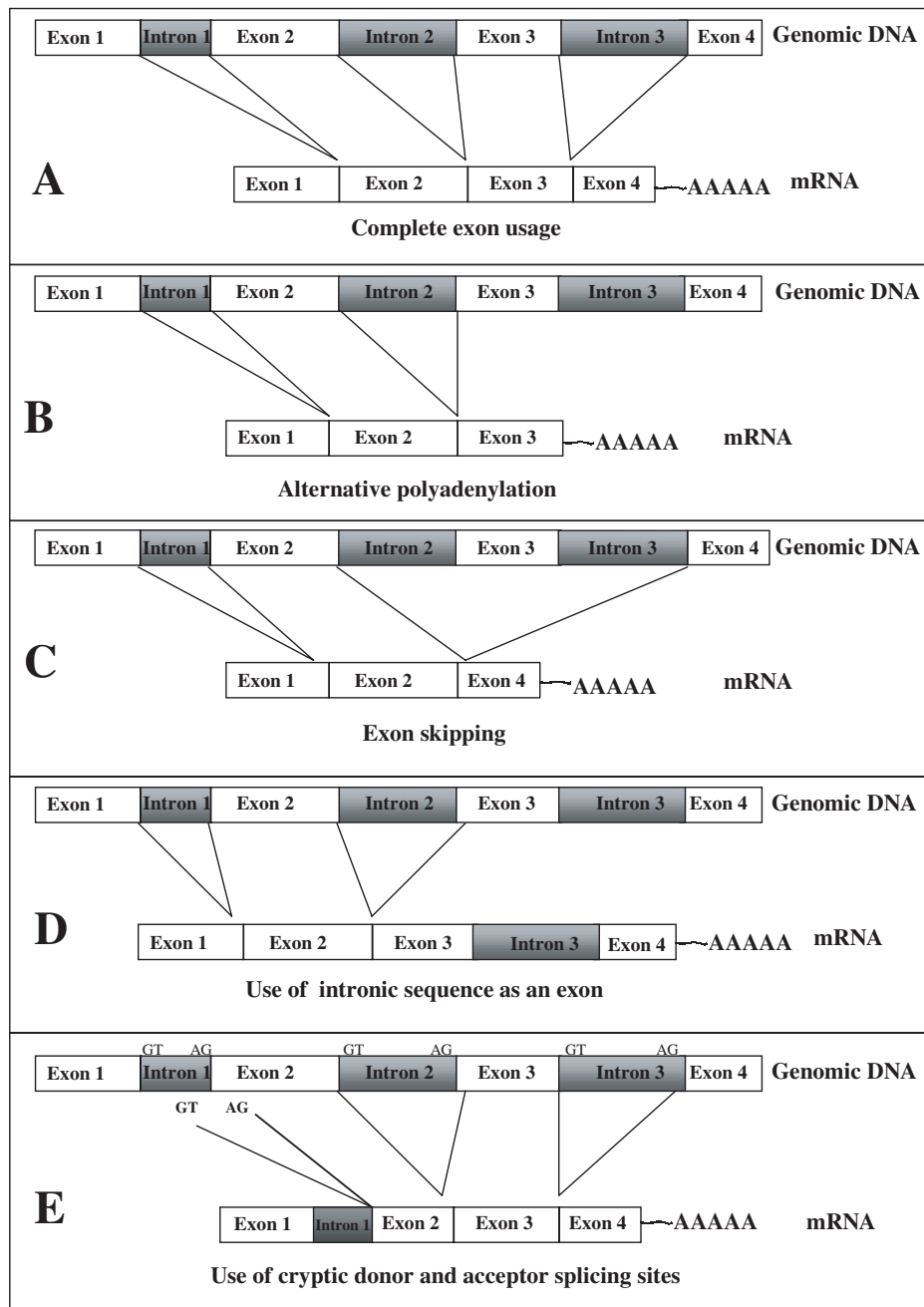


Fig. 1. The RNAs of some genes follow patterns of alternative splicing, whereby a single gene gives rise to more than one mRNA sequence. The majority of genes are transcribed into RNA giving rise to a single type of transcribed mRNA. Alterations in the polyadenylation site allow production of multiple mRNAs. Exons can be substituted, added or deleted. Introns that are normally excised can be retained in the mRNA. The positions of either 5' or 3' splice sites can shift to make exons longer or shorter.

estimate that 35% of human genes exhibit alternative spliced products [7]. The same rate was observed by Hanke et al. [8] in a set of 475 human proteins that were aligned to the human EST databases. It is widely thought, however, that this value is probably an under-

estimate since the transcript sequence information is derived from a limited number of tissues and developmental stages and as yet covers only a fraction of the human transcriptome. Lander et al. [9], for example, have found that 59% of all genes mapped on chromo-

some 22 have more than one splicing variant. The rate of alternative splicing in human genes already detected, however, seems to be considerably higher than in *C. elegans* where approximately 20% of all genes have at least 2 splicing variants. It is quite common to find human genes with dozens of splicing variants, including the neurexins, N-cadherins and potassium channels. Thus the current estimate of ~ 30,000 human genes may translate to hundreds of thousands of proteins.

3. Alternative splicing in cancer cells

Many cancer-associated genes are alternatively spliced. Loss of fidelity or variation of the splicing process occurs during tumor progression and may play a major role in tumorigenesis [10–14]. In addition, controlled switching to specific splicing alternatives may also occur. Transcript variants that only occur in tumors are both potential novel drug targets as well as potential diagnostic markers. In addition, their detailed analysis may prove crucial to our eventual understanding of the phenomena of malignancy. Probably the best characterized of known cancer associated genes that exhibit cancer associated alternative splicing are CD44 and WT1. These two genes represent two quite different situations. CD44 is a gene, similar to the *Drosophila* DSCAM gene that is apparently designed to exhibit considerable variability and has a large number of variably used exons. WT1 on the other hand has far fewer alternatives but which are of extreme importance to the malignant process.

3.1. CD44

The CD44 gene generates a family of molecules consisting of many isoforms [15]. CD44 proteins are single chain molecules comprising an N-terminal extracellular domain, a membrane proximal region, a transmembrane domain, and a cytoplasmic tail. The extracellular domain is glycosylated. CD44 is the principal hyaluronic acid (HA) receptor, although the molecule can bind other ligands, in some cases with low affinity. The CD44 proteins bind extracellular matrix glycoproteins, such as collagens and fibronectin. The CD44 gene has only been detected in higher organisms and the amino acid sequence of most of the molecule is highly conserved between mammalian species.

The molecular diversity of this glycoprotein is generated by both post-translational modification and differential exon utilization (Fig. 2). CD44 is encoded by

a single gene composed of 20 exons, located on the short arm of chromosome 11, spanning approximately 50 kb of human DNA [16]. The first 5 exons coding for the extracellular domain are designated the 5' constant region, whereas the next 10 exons are subjected to alternative splicing. This generates a variable region containing different exon combinations [15]. Variable region exons are designated V1 to V10. Exons 16 and 17 are the first two constant exons of the 3' constant region and they, together with part of exon 5, encode the membrane proximal region of the extracellular domain (with optional inclusion of variant exons). The next domain is the hydrophobic transmembrane region, which is encoded by exon 18 of the 3' constant region. The cytoplasmic domain is also subjected to alternative splicing. Differential utilization of exons 19 and 20 generates the short version (3 amino acids) and the long version (70 amino acids) of the cytoplasmic tail, respectively. The first 3 amino acids, common to both tails, are encoded by exon 18. The DNA sequence of exon 19 carries a long poly A+T tract, possibly causing instability in the mRNA of the short version. The additional amino acids of the long cytoplasmic domain are encoded by exon 20. The long version of the cytoplasmic tail is much more abundant than the shorter version [15]. The most abundant version of CD44 is the standard lacks the entire variable region, with exon 5 of the constant 5' region being directly spliced to exon 16 of the 3' constant region [15,17]. Individual cells can simultaneously express different isoforms [15].

The major physiological role of CD44 is to maintain organ and tissue structure via cell-cell and cell-matrix adhesion, but other isoforms can also participate in cell traffic, lymph node homing, presentation of chemokines and growth factors to traveling cells and transmission of growth signals [18].

The physiological functions of CD44 indicate that the molecule has characteristics that are consistent with it playing a role in the metastatic spread of tumors. Many studies have detailed the pattern of CD44 splicing and the transcript abundance in tumors. It has been found that changes in CD44 expression (mainly up-regulation, occasionally down-regulation, and frequently alteration in the pattern of isoforms expressed) are associated with a wide variety of cancers and the degree to which they spread. This is not universal, however, in some types of cancers, the CD44 pattern remains unchanged. Most importantly, the expression of CD44 has been shown to correlate with the progression and prognosis of some malignant tumors.

Recent studies have shown that CD44 is involved in two of the three steps of the invasive cascade: adhesion

sive colorectal cancer, the loss of standard-type CD44 expression in the invaded area is a sensitive marker for metastasis to the lymph nodes [24].

The expression of CD44 isoforms has also been evaluated in breast infiltrating lobular carcinomas in a panel of 39 tumors. The expression of membranous and cytoplasmic CD44s, V3, V5, V6, V7 and V3-10 was analyzed in the infiltrating cells by immunohistochemical staining. The protein positive tumors showed membranous and/or cytoplasmic staining with all antibodies used except for CD44V7, which only displayed cytoplasmic staining. Cytoplasmic expression of CD44V3 and membranous expression of V6 were significantly associated with alveolar, classical/alveolar carcinomas and mucinous/alveolar carcinomas. Furthermore, in alveolar, classical/alveolar and mucinous/alveolar carcinomas, cytoplasmic staining of CD44V5 was correlated with lymph node negative patients, whereas membranous V5 was correlated with lymph positive patients. In classical, classical/trabecular and trabecular carcinomas expression of membranous CD44s was significantly correlated with lymph node status [25]. The serum levels of different soluble CD44 molecules (CD44 standard form and CD 44 splice variant V6) were measured with an enzyme immunoassay method in venous blood samples preoperatively collected from 100 patients with invasive breast carcinoma. Preoperative serum soluble CD44 V6 was found to be closely related to distant metastases and TNM staging, indicating that CD44 V6 may have prognostic value in breast carcinoma.

It is becoming clear that CD44 variants may be useful as diagnostic or prognostic markers in some human malignant diseases. However, the data are conflicting, and further studies are needed to establish the prognostic value of CD44 and its variant isoforms. Furthermore, the precise function of CD44 in the metastatic process and the degree of involvement in human malignancies has yet to be established. Nevertheless, the studies cited above provide one of the most advanced instances of the systematic study of the association of particular alternatively spliced isoforms of a protein and malignancy and provides a proof of principle that alternatively spliced transcripts can act as cancer markers.

3.2. *WT-1*

Wilms tumor (WT) or nephroblastoma is a pediatric kidney cancer arising from pluripotent embryonic renal precursors. Multiple genetic loci have been linked

to Wilms tumorigenesis; positional cloning strategies have led to the identification of the WT1 tumor suppressor gene at chromosome 11p13 [26]. WT1 encodes a zinc finger transcription factor that is inactivated in the germline of children with genetic predisposition to Wilms tumor and in a subset of sporadic cancers. When present in the germline, specific heterozygous dominant-negative mutations are associated with severe abnormalities of renal and sexual differentiation, pointing to an essential role of WT1 in normal genitourinary development [27].

WT1 encodes a DNA binding protein that is thought to act as a transcriptional regulator. Exons 1–6 of WT1 encode domains involved in transcriptional regulation, dimerization, and possibly RNA recognition, whereas exons 7–10 encode four zinc fingers of the DNA-binding domain. Four isoforms of WT1 are formed by alternate RNA splicing [28] (Fig. 3), but in total, twenty-four potential protein isoforms may be synthesized due to contribution of two alternative splicing regions corresponding to the whole of exon 5 (17 amino acids) and to the three last codons of exon 9 (KTS), respectively [28], a site of RNA editing at codon 281 in exon 6 (a C to T transition producing a leucine to proline substitution [29], a non-AUG initiation codon resulting in WT1 proteins with a higher molecular weight [30] and an internal AUG initiation codon resulting in WT1 proteins with a lower molecular weight [31]. Biochemical and genetic evidence is accumulating that the WT1(–KTS) and WT1(+KTS) isoforms have different functions. WT1(–KTS) behaves as a transcription factor and *in vitro* can regulate several genes expressed during kidney development, including IGF2, PDGFA, EGFR, PAX-2, and WT1 [32]. However, the precise physiological and functional significance of this regulation is still unknown. The most frequent splice variants include the additional 17 aminoacids inserted N-terminal to the first zinc finger through the inclusion of exon 5. Insertion of the KTS tripeptide has a profound effect on both the DNA-binding affinity and the specificity of WT1. The WT1-KTS isoform binds to a 9-bp early growth response protein (EGR-1) consensus site with high affinity whereas the +KTS splice variant binds to the same site 10- to 20-fold more weakly [33]. Moreover, the different KTS+/- WT1 isoforms localize to distinct compartments in the nucleus suggesting that these two different forms of the protein have different functions. WT1-KTS co localizes with other transcription factors, whereas the more abundant +KTS isoform co localizes and is physically associated with splice factors where it potentially functions through

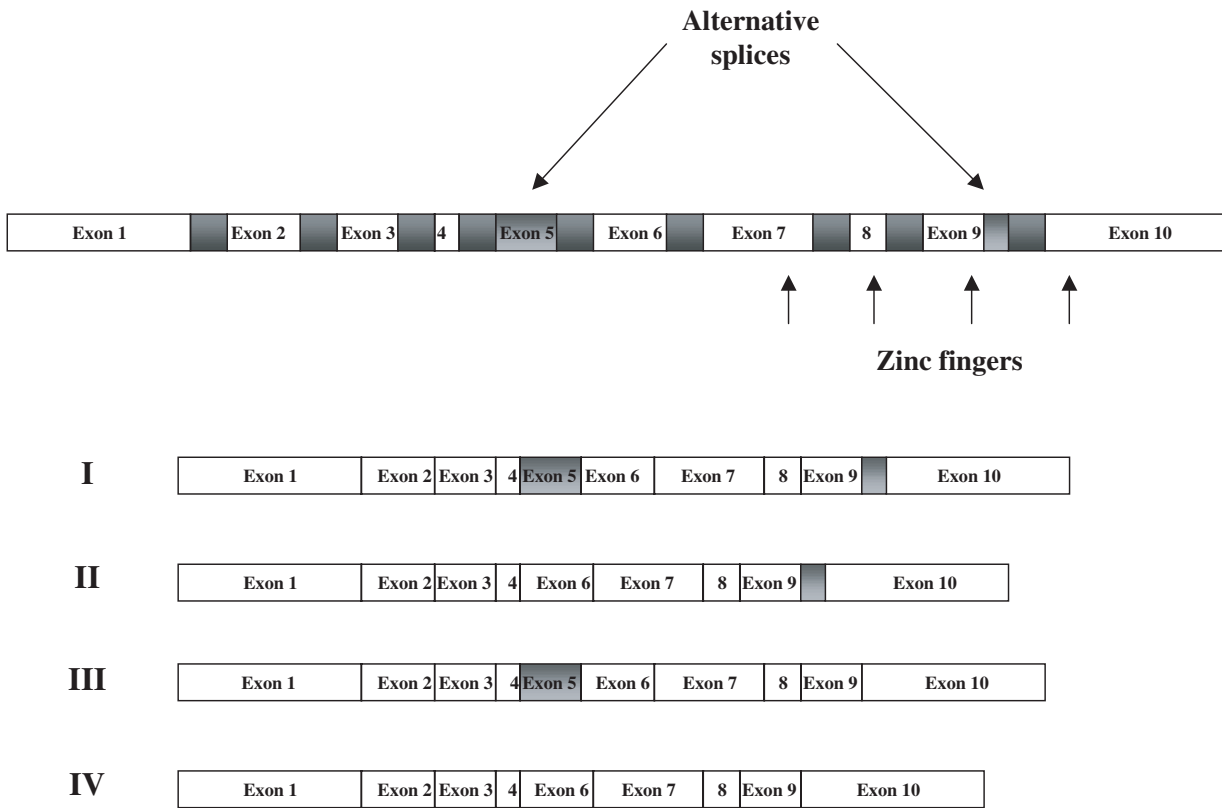


Fig. 3. Representation of WT1 genomic organization and alternatively spliced transcripts. The exons of WT1 (1–10) are shown in open boxes, while the introns and alternative splice sites are shaded. The carboxyl terminus of this protein contains four zinc finger domains. Four different transcripts (I to IV) are the result of alternative splicing. The most prevalent is isoform I.

RNA binding [34]. Thus, the presence or absence of the KTS insert in the third linker modulates both the DNA-binding affinity and the functional distribution of WT1 within the nucleus. Indeed, it has been suggested that differences in DNA-binding affinity of the +KTS and KTS splice variants might dictate the pattern of nuclear localization, with the tighter-binding KTS isoform being preferentially compartmentalized with the DNA. The balance between isoforms with and without the 17-amino acid insertion seems to affect the regulation of proliferation, differentiation, and apoptosis and the prevention of tumor formation [35,36].

In Wilms tumors, changes in the WT1 expression were found in 90% of unilateral unifocal WT cases, with 63% showing splicing alterations. Disruption of exon 5 splicing was the most frequent alteration, but alteration of exon 9-KTS splicing with an increase in the amount of isoforms with the KTS domain has also been observed in some tumors [12]. These results raise the possibility that regulation of splicing is an important factor in the development of the genitourinary system,

and that tumors may arise through aberrant splicing. WT1 isoform imbalance may be involved in various types of cancer because it has also been reported in breast tumors [37]. The distinct functional properties of WT1 isoforms and tumor-associated variants may shed light into the link between normal organ-specific differentiation and malignancy.

The WT1 gene contrasts starkly with CD44, since very discrete and limited transcript variability appears to have a profound effect on protein function. Given the very high percentage of human genes that exhibit alternative splicing of this type, it is clear that a huge, and as yet very poorly explored, wealth of biological information exists that could profoundly alter our understanding of cancer and our ability to detect and treat this devastating disease.

4. Conclusion

These two examples provide the first glimpse of the enormous potential importance of alternative splicing

to the understanding detection and treatment of human cancer. Alternate transcripts lend themselves well to laboratory-based detection. For example, one can envisage microarrays specifically composed of exons that we know to be alternatively spliced and which could prove significantly more powerful than other microarray types for cancer diagnosis and prognosis. On a more specific basis, individual transcript variants of particular relevance to the disease can be readily detected and quantified by established PCR technologies. Furthermore, since the alternative transcripts lead to altered protein structure, it should be possible to generate monoclonal antibodies that distinguish between isoforms. These will be applicable both in immunohistochemical analysis and also for ELISA determinations. Given the relatively small number of genes in the human genome there is widespread expectation that the complexity of transcript variants may prove to be in the key in understanding the functioning of the human body. We predict that they might be equally crucial in the development of complex genetic diseases such as cancer. The systematic exploration of this possibility over the current decade may thus be one of the most important routes to the development of the next generation of cancer markers.

References

- [1] A.J. Lopez, Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation, *Annu Rev Genet* **32** (1998), 279–305.
- [2] C.W. Smith and J. Valcarcel, Alternative pre-mRNA splicing: the logic of combinatorial control, *Trends Biochem Sci* **25** (2000), 381–388.
- [3] D.L. Black, Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology, *Cell* **103** (2000), 367–370.
- [4] B.R. Graveley, Alternative splicing: increasing diversity in the proteomic world, *Trends Genet* **17** (2001), 100–107.
- [5] D. Schmucker, J.C. Clemens and H. Shu et al., *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity, *Cell* **101** (2000), 671–684.
- [6] R. Tacke and J.L. Manley, Determinants of SR protein specificity, *Curr Opin Cell Biol* **11**(3) (1999), 358–362.
- [7] L. Croft, S. Schandorff, F. Clark, K. Burrage, P. Arctander and J.S. Mattick, ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome, *Nat Genet* **24** (2000), 340–341.
- [8] J. Hanke, D. Brett and I. Zastrow et al., Alternative splicing of human genes: more the rule than the exception? *Trends Genet* **15** (1999), 389–390.
- [9] E.S. Lander, L.M. Linton and B. Birren et al., Initial sequencing and analysis of the human genome, *Nature* **409** (2001), 860–921.
- [10] B. Kwabi-Addo, F. Ropiquet, D. Giri and M. Ittmann, Alternative splicing of fibroblast growth factor receptors in human prostate cancer, *Prostate* **46** (2001), 163–172.
- [11] T.I. Orban and E. Olah, Expression profiles of BRCA1 splice variants in asynchronous and inG1/S synchronized tumor cell lines, *Biochem Biophys Res Commun* **280** (2001), 32–38.
- [12] D. Baudry, M. Hamelin and M.O. Cabanis et al., WT1 splicing alterations in Wilms' tumors, *Clin Cancer Res* **6** (2000), 3957–3965.
- [13] M. Mixon, F. Kittrell and D. Medina, Splice variant expression of CD44 in patients with breast and ovarian cancer, *Oncol Rep* **8** (2001), 145–151.
- [14] M. Hori, J. Shimazaki, S. Inagawa, M. Itabashi and M. Hori, Alternatively spliced MDM2 transcripts in human breast cancer in relation to tumor necrosis and lymph node involvement, *Pathol Int* **50** (2000), 786–792.
- [15] D. Naor, R.V. Sionov and D. Ish-Shalom, CD44: structure, function, and association with the malignant process, *Adv Cancer Res* **71** (1997), 241–319.
- [16] G.R. Screaton, M.V. Bell, D.G. Jackson, F.B. Cornelis, U. Gerth and J.I. Bell, Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons, *Proc Natl Acad Sci USA* **89** (1992), 12160–12164.
- [17] I. Stamenkovic, A. Aruffo, M. Amiot and B. Seed, The hematopoietic and epithelial forms of CD44 are distinct polypeptides with different adhesion potentials for hyaluronate-bearing cells, *EMBO J* **10** (1991), 343–348.
- [18] R.J. Sneath and D.C. Mangham, The normal structure and function of CD44 and its role in neoplasia, *Mol Pathol* **51** (1998), 191–200.
- [19] A. Herrera-Gayol and S. Jothy, Adhesion proteins in the biology of breast cancer: contribution of CD44, *Exp Mol Pathol* **66** (1999), 149–156.
- [20] J. Lesley, R. Hyman, N. English, J.B. Catterall and G.A. Turner, CD44 in inflammation and metastasis, *Glycoconj J* **14** (1997), 611–622.
- [21] P. Herrlich, H. Morrison and J. Sleeman et al., CD44 acts both as a growth- and invasiveness-promoting molecule and as a tumor-suppressing cofactor, *Ann N Y Acad Sci* **910** (2000), 106–118.
- [22] N. Harada, T. Mizoi and M. Kinouchi et al., Introduction of antisense CD44S CDNA down-regulates expression of over-all CD44 isoforms and inhibits tumor growth and metastasis in highly metastatic colon carcinoma cells, *Int J Cancer* **91** (2001), 67–75.
- [23] S. Aaltomaa, P. Lipponen, M. Ala-Opas and V.M. Kosma, Expression and Prognostic Value of CD44 Standard and Variant v3 and v6 Isoforms in Prostate Cancer, *Eur Urol* **39** (2001), 138–144.
- [24] T. Asao, J. Nakamura and Y. Shitara et al., Loss of standard type of CD44 expression in invaded area as a good indicator of lymph-node metastasis in colorectal carcinoma, *Dis Colon Rectum* **43** (2000), 1250–1254.
- [25] H.S. Berner and J.M. Nesland, Expression of CD44 isoforms in infiltrating lobular carcinoma of the breast, *Breast Cancer Res Treat* **65** (2001), 23–29.
- [26] K.M. Call, T. Glaser and C.Y. Ito et al., Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus, *Cell* **60** (1990), 509–520.
- [27] S.B. Lee and D.A. Haber, Wilms tumor and the wt1 gene, *Exp Cell Res* **264** (2001), 74–99.
- [28] D.A. Haber, R.L. Sohn, A.J. Buckler, J. Pelletier, K.M. Call and D.E. Housman, Alternative splicing and genomic structure of the Wilms tumor gene WT1, *Proc Natl Acad Sci USA* **88** (1991), 9618–9622.

- [29] P.M. Sharma, M. Bowman, S.L. Madden, F.J. 3rd Rauscher and S. Sukumar, RNA editing in the Wilms' tumor susceptibility gene WT1, *Genes Dev* **8** (1994), 720–731.
- [30] W. Bruening and J.A. Pelletier, non-AUG translational initiation event generates novel WT1 isoforms, *J Biol Chem* **271** (1996), 8646–8654.
- [31] V. Scharnhorst, P. Dekker, A.J. van der Eb and A.G. Jochemsen, Internal translation initiation generates novel WT1 protein isoforms with distinct biological properties, *J Biol Chem* **274** (1999), 23456–23462.
- [32] A. Menke, L. McInnes, N.D. Hastie and A. Schedl, The Wilms' tumor suppressor WT1: approaches to gene function, *Kidney Int* **53** (1998), 1512–1518.
- [33] W.A. Bickmore, K. Oghene, M.H. Little, A. Seawright, V. van Heyningen and N.D. Hastie, Modulation of DNA binding specificity by alternative splicing of the Wilms tumor wt1 gene transcript, *Science* **257** (1992), 235–237.
- [34] S.H. Larsson, J.P. Charlieu and K. Miyagawa et al., Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing, *Cell* **81** (1995), 391–401.
- [35] S.M. Hewitt and G.F. Saunders, Differentially spliced exon 5 of the Wilms' tumor gene WT1 modifies gene function, *Anti-cancer Res* **16** (1996), 621–626.
- [36] C. Englert, X. Hou and S. Maheswaran et al., WT1 suppresses synthesis of the epidermal growth factor receptor and induces apoptosis, *EMBO J* **14** (1995), 4662–4675.
- [37] G.B. Silberstein, K. Van Horn, P. Strickland, C.T. Jr. Roberts and C.W. Daniel, Altered expression of the WT1 wilms tumor suppressor gene in human breast cancer, *Proc Natl Acad Sci USA* **94** (1997), 8132–8137.