1. Introduction

In this review, we discuss the merit of splicing isoforms as a source of biomarkers for ovarian cancer with a special focus on features that distinguish splice variants from global gene expression based markers. Key examples demonstrating the usefulness of alternative splicing (AS) as markers of ovarian cancer are described.

Ovarian cancer is a low incidence cancer with high mortality rate [1]. The asymptomatic nature of this cancer and the late stage diagnosis of most tumors are the reasons for ineffective surgery and chemotherapy [2]. In this sense, intensive research aims at increasing overall patient survival and quality of life by providing biomarkers for 1) early detection and 2) prediction of chemotherapy response and/or suggestion of alternative strategies. CA-125 is a glycoprotein that is usually expressed in a variety of epithelial cells and its serum level rise up in advance ovarian cancer [3]. However, its use as an early detection marker or as a tool to screen the general population has not been approved so far [4,5]. CA-125 level is helpful in treatment-decision making but do not retain the capacity to improve overall survival and quality of life [6,7]. Clearly, there is still great need for biomarkers or combination of biomarkers that could positively identify early ovarian cancer lesions with great certainty or increase patients’ survival.

Genome-wide mRNA profiling presents an opportunity to rapidly identify RNA markers. Microarray platform has been applied in numerous occasions to provide gene expression signature correlating prognosis or indicative of chemotherapy response (review in [8]). However, the nature of the platform used to carry the experiments and the analysis methods and sample sets makes inter-laboratory comparison very difficult and finding reliable mark-
ers complicated. Indeed, a meta-analysis regrouping 829 samples fails to demonstrate the predictive power of 16 individual gene expression signatures [9]. Consequently, very few microarray markers reached the clinic. In contrast, high-throughput protein signature based on mass spectrometry platform appears to have much more overlap in the peaks found by different experimental studies [10]. However, the pace by which protein biomarkers are translated into clinical setting is relatively slow [11]. Clearly, there is a need for novel methodology to discover ovarian cancer biomarkers that can yield reliable results and produce tests that could be quickly integrated in normal clinical setting. In this chapter, we discuss the potential of splice variant annotations as a tool for the discovery of ovarian cancer markers and discuss the challenges and promises of this hidden mine.

2. Pre-mRNA splicing mechanism and regulation

Transcription of messenger RNA (mRNA) is the first step of converting the DNA code into functional proteins. This process was often seen as a linear cascade of events that include mRNA capping [12], splicing [13], polyadenylation [14], export to the cytoplasm [15] and translation [16] to produce a single protein. However, in reality a single pre-mRNA can produce many mRNAs through the process of AS and this in turn lead to the production of several proteins from a single gene. Splicing is the process by which the protein coding exons (typically hundreds of nucleotides in length) are joined together after the removal of large non-coding introns (typically thousands of nucleotides in length) to form the coding sequence. In some genes, this process leads to one outcome and thus named constitutive (Fig. 1.) but in most cases it leads to more than one outcome and thus called alternative (Fig. 2). Both processes are mediated by the spliceosome, specialized machinery that recognizes consensus RNA sequences [13,17]. The spliceosome component U1snRNP binds the 5’ splice site (5’ss), the splicing Factor 1 (SF1) binds the branch point site (BPS) adenine, U2 auxiliary factor 65 kDa subunit (U2AF65) binds the poly-pyrimidine tract (PPT) and U2 auxiliary factor 35 kDa subunits (U2AF35) binds the 3’ splice site (3’ss) (Fig. 1A). The last two component are further replaced by U2snRNP and following complex base-paring rearrangements and RNA-protein interactions involving hundreds of protein, the spliced mRNA, the intron by-product and the spliceosome component are release [17]. Chemically speaking, the splicing reaction proceeds in two trans-esterification steps (Fig. 1B). The first step involved the attack by the 2’ hydroxyl of the branch point adenine on the phosphate at the 5’ss, releasing at the same time the 3’end of the mRNA. The second step involved the attack by one of the hydroxyl of the terminal phosphate on the phosphate at the 3’ss, liberating the intron in the form of a lariat. This cycle of spliceosome assembly/disassembly is repeated for every intron of a gene on the nascent RNA transcript [18].

When the splice site for some exons become weak or introns with suboptimal sequence exist splicing may become less accurate and may depend on the factors that influence the splicing of competing exons and consequently produce mRNA versions with different exon pairs. AS affected the majority of multi-exons genes and is believed to be the principal driver of proteome diversity [19]. As illustrated in figure 2, two 5’ss can compete for a single 3’ss or inversely, two 3’ss can compete for a single 5’ss. These type of alternative splicing events (ASEs) are referred to alternative 5’ (alt5’, Fig. 2A) and alternative 3’ (alt3’ Fig. 2B), respectively. The most frequent
type of ASE in human is the full skipping of an exon (cassette exon, Fig. 2C). Some exons are also skipped as a bloc (multiple cassette exon, Fig. 2D) or mutually exclusive (Fig. 2E). AS could also be coupled to others regulatory mechanisms such as polyadenylation [14] (Fig. 2F). In this case, the resulting mRNA exhibits a different 3’ untranslated region (UTR), which is further subjected to different regulation by small non-coding RNA (e.g. microRNA). In about 1 out of 3 cases, AS decision introduces a sequence containing premature stop codon [20]. In these cases, the resulting mRNA is flagged to be degraded by the non-sense mediated decay machinery creating an efficient mechanism that control gene expression post-transcriptionally [21] (Fig. 2G). In some cases, ASE occurs outside the coding region and influence regulatory sequence in the UTR [22]. These different forms of splicing isoforms should not be confused with those generated by alternative transcription start site where a single gene might transcribe from different promoters (Fig. 2H). In this case, and unlike alternative splicing, there is little chance that the isoform will have different protein sequence unless a new protein-coding exon is added in frame for translation initiation.

Figure 1. Constitutive splicing. A) Consensus splicing sequences. The 5’ss, BPS, PPT and 3’ss are represented and are binds by U1 snRNP, SF1, U2AF65 and U2AF35, respectively. B) Splicing reaction. The first step involved the attack by the 2’ hydroxyl of the branch point adenine on the phosphate at the 5’ss, releasing at the same time the 3’end of the mRNA. The second step involved the attack by one of the hydroxyl of the terminal phosphate on the phosphate at the 3’ss, liberating the intron in the form of a lariat.

ASEs are normally associated with low sequence conservation near the splice site and instead are usually linked to RNA binding motifs that may enhance or repress exon inclusion [23,24]. Motifs that enhance exon inclusion often recruit splicing factors like the SR protein family, which in turn interact with the spliceosome via an arginine serine rich domain to increase weak 5’ss and 3’ss recognition [25] (Fig. 3A). On the other hand, splicing motifs that promotes exon exclusion by binding members of the hnRNP family oligomerized through exon [26], block UsnRNA recruitment [27] or loop out the alternative exon [28] (Fig. 3B to D). Similarly, sequence motifs in intron may bind to SR or hnRNP proteins to influence splicing, but in this case, the SR proteins results in exon exclusion and hnRNP in exon inclusion. This is most likely because hnRNPs define intronic region and SR protein define exons location. Usually, these different enhancers and repressor protein families work together to define the final outcome of any ASEs (Fig. 3E) [29,30]. One of the most conserved intronic motif downstream of alternative exons is the UGCAUG motif [31,32], which bind the tissue-specific splicing factors family RBFOX. In general it is suggested that tissues specific splicing factors favor exon inclusion when bound to introns downstream of alternative exons and exclusion when bound upstream. This rule is beginning to be appreciated for several splicing factor such as Celf [33], epithelial-specific regulatory protein [34], Nova [35] and RBFOX [36] (Fig. 3F).
3. The advantage of alternative splicing as a source of ovarian cancer biomarker

Analysis of the ovarian cancer proteome using mass spectrometry is undoubtedly the most direct approach for the identification of biomarkers that could be readily implemented in the clinic. However, the difficulty generating specific antibodies for the large number of potential markers generated via this approach makes marker validation very difficult. In contrast, the validation of nucleic acid markers generated through microarray or deep-sequencing screen is fairly simple and is often achieved by polymerase chain reaction (PCR) [34, 37]. Furthermore, the function of these potential markers can easily be verified through the knockdown of gene expression using RNA interference (RNAi) strategies [38]. However, scoring global changes in gene expression as markers for ovarian cancer limits the assay to ~25000 genes in the genome, while it is estimate that the human cells contains at least >100 000 proteins. This limitation is no longer an issue when we consider the expression of specific splice variants, the number of which equal or exceeds the number of cellular proteins [39]. In addition, it is much easier to predict the function of an alternative splice variant than predicting the function of a peptide marker. For example, while the role of the well established markers CA-125 remain unclear after 25 years of research [40], one could easily predict the function of a marker by the protein domain eliminated or included through AS as is the case of the tyrosine kinase SYK. In this case, exon skipping remove a nuclear localization domain leading to the accumulation of protein in the...
cytoplasm, elegantly explaining the lost of nuclear function associated with cancer [38]. Predicting the impact of AS is particularly attractive for biomarkers development when the alternative exon encodes a plasma transmembrane domain or an extracellular protease cleavage site [41]. In these cases, one would be able to predict whether the cancer associated marker leads to an increase or decrease in the secretion of membrane anchored protein, an information that is difficult to obtain using global gene expression profiles.

4. The challenges of detecting splicing isoforms

Examples of alternatively spliced genes are steadily accumulating in the literature for more than 20 years and the discovery rate was greatly accelerated by recent technological advances like transcriptome sequencing techniques. Indeed, while early estimation of alternatively spliced genes based on Northern-Blots and endpoint RT-PCR were around 5% of the human genome, transcriptome sequencing revealed ASEs in 95% of the human genes with multiple introns [39]. Different techniques have different capacity to illustrate the number of ASEs (see Table 1) and detecting splice variants remained difficult to detect for many years, which explains the reason they are not regularly considered as a source of biomarkers by most clinicians.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Throughput</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern-Blot</td>
<td>Low</td>
<td>No amplification</td>
<td>Labour-intense</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Several isoforms can be detected in a single sample</td>
<td>Large amount of RNA needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Restricted by gel resolution</td>
</tr>
<tr>
<td>Endpoint RT-PCR</td>
<td>Low to medium</td>
<td>Considered as the gold standard for validation</td>
<td>Labour-intensive if polyacrylamide gels are used to separate PCR products</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Results easy to analyze</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Throughput is enhanced when coupled to capillary electrophoresis</td>
<td>Low quantitative range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Restricted by gel resolution</td>
</tr>
<tr>
<td>Real-time PCR</td>
<td>Low to medium</td>
<td>Provided already validated data</td>
<td>Custom primer design often needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large quantitative range</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accurate data in fixed tissues</td>
<td></td>
</tr>
<tr>
<td>Microarray</td>
<td>medium</td>
<td>Some array are commercially available</td>
<td>Complex analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Results need PCR validation</td>
</tr>
<tr>
<td>Next generation Sequencing</td>
<td>Medium to High</td>
<td>Independent of genome annotation</td>
<td>High cost prevent the use of biological replicates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Discovery of novel splicing isoforms)</td>
<td>Long multi-step procedure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Complex analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Results need PCR validation</td>
</tr>
</tbody>
</table>

Table 1. Techniques used for the detection of alternative splicing
Back in the 1980's, splicing isoforms were mainly detected by Northern-Blot, which separate transcripts by size [42] and estimate relative mRNA abundance using internal controls. However, this method is difficult to adopt in a clinical setting and require a large amount of RNA (µg), which is difficult to obtain from clinical samples. Later, the discovery of reverse transcription and PCR amplification greatly facilitated the detection of splice variants [43]. Splicing isoform amplification is achieved by using PCR primers that are designed to hybridize to constitutive exons flanking the ASE of interest (Fig. 4A). The products are separated in agarose gels or capillary gel electrophoresis [44] and the ratio of the long and short isoform quantified and presented as ψ (percent of splicing index): the molarity of the long over the sum of the long and the short isoforms (Fig. 4A). Even if competitive PCR reaction are limited to a narrow range [43,45], endpoint PCR is still the preferred technique to detect splicing isoforms due to the ease of use and low cost of the experiments.

The gold standard for the mRNA quantification is real-time PCR [46], which unlike standard endpoint PCR, detects the amount of products accumulating after each cycle of amplification and permits accurate comparison of different samples. This type of PCR requires the use of fluorescent probes [47] or dyes [48] that permit detection by specialized sensors. Despite the accuracy of this detection method it is rarely used for the detection of splice variants due to difficulty in achieving isoform specific amplification [43,49]. Primers required for the amplification of the short isoform need to bind to a short unique sequence created by the exon-exon junction, which severely restrict the design (Fig. 4B). However, systematic evaluation of isoform specific design parameters and the availability of new algorithms for primer selection greatly facilitated the detection of ASEs from any species [49]. Indeed, universal PCR conditions and ease of primer design makes real-time PCR reaches the point where it can compete with high-throughput detection methods like microarray in term of ASE coverage [49].

Microarray as a method for genome-wide expression profiling was discovered in 1995 [50], but the use of this method to detect splice variants was reported only in 2003 [51]. It took 8 years to develop methods that could distinguish between the hybridization patterns of two closely related transcripts and develop chips with high enough density to accommodate the thousands of splicing isoforms [51] (Fig. 4C) Early attempts to extract splicing pattern from expression microarrays generated high false positive rate [52]. Therefore, strategies where developed to probe exon-exon junction (junction array) [51]. In this case, alternative exons are defined by very low or very high signals emanating from two consecutive splice junctions [51]. Another popular strategy is to use exonic probe in addition to exon-exon junction probe (exon/junction array) [53]. In every case, the high similarity of exon-exon junction to favor non-specific hybridization and in some analysis procedures the information is restricted to splicing isoform “detection” rather than true quantification [54]. The most successful quantification of splicing isoform by microarray was achieved by relying solely on exonic probe [54-56]. However, the success of this method was limited by its dependence on a small set of pre-selected splice variants [53,57]. To allow the discovery of new splicing isoforms, a fourth strategy that consider all putative exons (tiling array) was developed [58]. However, the high number of probe required for this methods restricted coverage to only a
small fraction of the genome. Not surprisingly, these difficulties hampered the application of this method for the study of ovarian cancer splicing isoforms. Indeed, to date there is no report of microarray based profiling of ovarian cancer splice variants.

In theory, the most promising approach for the detection of ovarian cancer splicing isoform is the transcriptome sequencing [39]. Next generation sequencing (NGS) technology provide massively parallel sequencing of nucleotidic sequences in miniaturized microsystem. Several platforms are commercially available and their unique technology are discussed elsewhere [59,60]. The specific application of mRNA quantification through sequencing (RNA-seq) was demonstrated for different cancer types (e.g. lung [61] and prostate [62]) but not ovarian cancer thus far. Encouraging development in the refinement of the analytical pipeline to allow accurate quantification of splicing isoforms was recently made [37,63,64]. However, the complexity of the analytical pipeline of sequencing data and the cost of the sequencing read necessary to detect splice variants will reduce the speed by which this technique is applied to the discovery of splicing dependent biomarkers (Fig. 4D). In addition, secondary techniques like PCR will still be needed to confirm and validate the accuracy of the data generated and confirm it in a large number of clinical sample. Indeed, the majority of the AS information in ovarian cancer are derived from PCR-based techniques (see Table 2 and 3).

**Figure 4.** Methodology for splicing isoforms detection. A) Endpoint PCR. B) Real-time PCR. C) Microarray. D) Next Generation Sequencing. Red arrows and lines refer to short isoform specific detection. Gray boxes = constitutive exons; white boxes = alternative exons.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Technique</th>
<th>Marker type and isoform description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA1/2</td>
<td>Endpoint RT-PCR</td>
<td>Numerous DNA mutation inactivated BRCA1/2 function (deletion of essential domain or protein truncation) through aberrant splicing.</td>
<td>[67]</td>
</tr>
<tr>
<td>RUNX1 (AML1)</td>
<td>Real-time RT-PCR</td>
<td>AML1b&lt;sup&gt;179-242&lt;/sup&gt; expression inversely correlates with overall survival.</td>
<td>[69]</td>
</tr>
<tr>
<td>KLF6</td>
<td>Real-time RT-PCR</td>
<td>The ratio of total to full length KLF6 expression correlates with ovarian tumor grade.</td>
<td>[72]</td>
</tr>
<tr>
<td>TP53</td>
<td>Real-time RT-PCR</td>
<td>p53δ expression is associated to impaired response to first line chemotherapy.</td>
<td>[75]</td>
</tr>
<tr>
<td>FBLN1</td>
<td>Endpoint RT-PCR</td>
<td>*The ratio of Fibulin C/D increase in ovarian tumors.</td>
<td>[77]</td>
</tr>
<tr>
<td>SPP1</td>
<td>Real-time RT-PCR</td>
<td>*Osteopontin-c is undetectable in normal ovaries and present in ovarian tumors.</td>
<td>[85]</td>
</tr>
</tbody>
</table>

*Potential serum marker

**Table 2.** Ovarian cancer splicing markers discovered through single gene analysis

<table>
<thead>
<tr>
<th>Number of ASEs</th>
<th>Technique</th>
<th>Marker type and splicing signature description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>48 ASEs</td>
<td>Endpoint RT-PCR</td>
<td>This signature distinguishes normal ovaries from ovarian tumors regardless of grade, stage or histotype (serous, mucinous, endometrioid, mix type)</td>
<td>[44]</td>
</tr>
<tr>
<td>288 ASEs</td>
<td>Endpoint RT-PCR</td>
<td>This signature distinguishes normal ovaries from serous high grade ovarian tumors.</td>
<td>[36]</td>
</tr>
<tr>
<td>8 ASEs</td>
<td>Real-time RT-PCR</td>
<td>This cancer epithelial signature (CES) distinguishes normal Fallopian tube epithelium tissues from ovarian epithelial cancer cells</td>
<td>[84]</td>
</tr>
<tr>
<td>10 ASEs</td>
<td>Real-time RT-PCR</td>
<td>This cancer stromal signature (CSS) distinguishes normal tissues from ovarian tumors independent of the epithelial content of the tissue compared</td>
<td>[84]</td>
</tr>
</tbody>
</table>

**Table 3.** Ovarian cancer splicing markers derived from high-throughput expression profiling.
5. Example of alternative splicing based ovarian cancer biomarkers

5.1. Gene specific discovery of splicing markers

PCR-based techniques of specific genes associated with ovarian cancer revealed a number of ovarian cancer associated splicing events. The most promising of these potential biomarkers for diagnosis, prognosis, chemoresistance and grade biomarkers are listed in table 2 and are further described in the text below.

**BRCA1/2.** The ability of DNA mutation in the hereditary gene BRCA1 and BRCA2 (BRCA1/2) to predict the risk of ovarian and breast cancer is known for decades [65,66]. However, in several instance, the clinical relevance of DNA mutation is unknown, making the clinical management difficult to establish properly. Furthermore, the role of these mutations in varying the splicing of BRCA genes was largely ignored despite the fact that any nucleotide changes in the splice site consensus sequences or in any A5 regulatory sequence could produce aberrant splicing isoforms. Recently, *in vitro* splicing assay of BRCA1/2 mutation [67], revealed that many cancer associated variants including those with unsuspected synonymous mutations have dramatic effect on splicing. Strikingly, six of the most frequent DNA variants representing 58.5% of BRCA1 families induced aberrant splicing profile [67]. These results, clearly demonstrate the importance of studying cancer associated splicing since it may in many case help explain mutation that cannot be associated with changes in protein sequence. However, the fidelity of splicing signature as a diagnostic marker as compared to DNA sequencing remains to be established.

**RUNX1.** Runt-related transcription factor 1 is a transcriptional regulator harboring a DNA-binding runt homology domain (RHD). Several layers of regulation (transcription, splicing and translation) fine-tune its tissue-specific expression [68]. RUNX1 is also known as Acute myeloid leukemia (AML) 1 and is often found as oncogenic fusion in leukemia. Nanjundan M and collaborators [69] report the fortuit discovery of a novel isoform that lack exon 6 as compared to isoform AML1b during classical cloning procedure. This novel isoform, subsequently named AML1b<sub>Del179-242</sub> was found to be the dominant isoform in the majority of the 42 ovarian tumors studied. Functionally speaking, skipping of exon 6 severely abrogated the transactivation potential of the resulting protein and inhibits its tumor suppressive functions. Interestingly, AML1b<sub>Del179-242</sub> level was either not different from normal cell line (lung cancer) or significantly decreased (breast cancer), suggesting that it may be an ovarian specific marker [69]. Nonetheless, AML1b<sub>Del179-242</sub> expression is inversely proportional to the survival rate of patient, suggesting is used as a prognosis marker [69]. As attractive as it sounds, AML1b<sub>Del179-242</sub> certainly represent an excellent potential target and marker but required further validation in larger cohort and by independent research group.

**KLF6.** Kruppel-like factor (KLF) 6 is a transcription factor from the well conserved KLF gene family implicated in differentiation, development and cell growth [70,71]. KLF6 is a suspected tumor suppressor gene in several epithelial cancer (see [72] and reference there in). In ovarian cancer, an increase in KLF6 isoforms was noted and correlates with the aggressivity of the tumor in tissues (grade). One of these isoform is produced by the use of the more dis-
tal alt5′ ss and produce a protein version lacking the characteristic zinc finger binding domain (KLF6 SV1) and act as a dominant-negative [72]. Although the technical challenges of amplifying specific KLF6 isoforms preclude pinpointing the KLF6 SV1 as the isoform that correlates with grade tumor, a series of in vitro and in vivo evidence making it likely [72]. It remains to be established whether or not the full length KLF6 over KLF6 SV1 isoform ratio could serve as a prognosis or even early marker.

TP53. The tumor suppressor gene TP53 is mutated in several solid tumors and in almost all of the serous ovarian tumors [73]. Mutations affecting splice site of TP53 are very common [74] and leads to a complex pattern of splicing isoforms that add up to the already complex picture of this gatekeeper. Indeed, the point mutation IVS9-2A>G destroys the splice acceptor and redirect the splicing to include exon 9c. The resulting p53δ protein isoform is truncated in the oligomerization domain and have a new stretch of 27 new residues. In a cohort of 245 ovarian samples, the expression of this isoform is significantly correlated with poor overall survival in multivariate analyses. Moreover, patient having tumor that express p53δ have a higher chance of early relapse after first-line chemotherapy [75]. Since isoform p53δ doesn’t correlate with the debulking status, it suggests that expression of p53δ impair platinum-based chemotherapy [75]. In respect to personalized medicine, it would be very interesting to sensitize these tumors by targeting p53δ. Thus, p53δ is not only a potential adverse prognosis marker but could also be a promising target for a subclass of ovarian tumors.

FBLN1. Fibulin (FBLN) 1 is an extracellular matrix (ECM)-associated protein produced by both stromal cells and epithelial cells [76]. Its presence near sites of epithelial cells locally invading stromal boundaries suggests its implication in cell adhesion/motility (see [77] and reference there in). Different C-terminal exons of the FBLN1 gene are alternatively spliced to generate four isoforms. In ovarian tumors, the ratio fibulin 1C / fibulin 1D is significantly increased compared to benign ovarian cystic sample [77]. Interestingly, the isoform ratio between normal and benign cyst is slightly increased in cystic samples (although not significant). It raises the hypothesis that the ratio fibulin 1C / fibulin 1D could potentially serves as an early diagnostic marker. Importantly, the sensitivity and specificity of fibulin splicing isoforms remains to be firmly established using a panel of normal and early lesion tissues and ultimately in patient’s serum.

SPP1. Osteopontin is a member of the small integrin-binding ligand, N-linked glycoprotein (SIBLING) family of proteins [78]. It is an important component of the ECM that is secreted by both cancer cells and stromal cells in the tumor microenvironment [79]. Osteopontin interacts with various integrin receptors [80,81] as well as the CD44 receptor [82] to activate the angiogenic switch or enhance cancer cell motility [79]. The level of osteopontin is elevated in patient’s plasma when compared to healthy controls by enzyme-linked immuno assay (ELISA) [83]. However, the specificity (80,4%) and sensitivity (80,4%) for the detection of early stage disease are not convincing [83]. These parameters could be increase if one takes advantages of AS. A recent report conducted by real-time PCR indicated that the isoform osteopontin c (excluding exon 4) is absent in normal or benign tissues but always present in ovarian cancer samples. This is supported by our own data from microdissected normal and cancerous ovarian cells indicating that the expression of osteopontin c comes specifically
from ovarian epithelial cancer cells [84]. Conditioned medium overexpressing osteopontin c stimulate proliferation of cancer cells more efficiently than either osteopontin a or b, and this effect is revert by specific antibodies against osteopontin c [85]. Based on these data, the biomarker capacity of osteopontin in patient’s blood need to be re-established using isoform-specific methodology. As the secretion of osteopontin might be an early event [86-88], it is tempting to speculate that osteopontin c could be an early marker.

5.2. Splicing markers generated through genome-wide expression profiling

The advent of splicing sensitive high-throughput technique opens the doors to monitor a large number of randomly selected ASEs rather than be limited to few candidate genes (see Table 3). The recent use of high-throughput RT-PCR by coupling PCR reaction in 384 wells plate to capillary gel electrophoresis in 96 well Caliper station dramatically increased the number of confirmed ovarian cancer associated splicing events. Initially, exon-exon junctions were systematically analyzed for a set of 600 cancer related genes in four different pools of normal and cancer ovarian samples. The resulting ASEs were subsequently validated using an independent set of 21 normal ovaries and 25 ovarian cancer samples, yielding 48 ASE markers [44]. Later on, a focus on a collection of 2168 highly curated ASEs (RefSeq NCBI build 36) subsequently yield 288 ASEs markers using roughly the same sample set [36]. The relatively high number of ASEs markers found coupled to the fact that several were related to the epithelial-mesenchymal transition raised the possibility that a large fraction of the discovered events might result from difference in the cell type compared (normal ovaries are largely composed of stromal cells where as ovarian tumors have a typical epithelial content around 75% [36]). This question was answered when 9 ovarian tumors were microdissected to isolated the RNA from stromal (tumor microenvironment) and epithelial cancer cells separately. A real-time PCR-based screening strategy coupled to an update version of RefSeq NCBI build 36 (3313 ASEs) yield a low but unambiguous set of cancer-specific splicing isoforms, the cancer epithelial signature (CES) [84]. Surprisingly, the tumor microenvironment appears to contain promising splicing isoforms RNA markers. Indeed, this cancer stromal signature (CSS) might be able to diagnosis early ovarian tumors as it clusters low malignant potential and low-grade tumors within normal ovaries and Fallopian tube samples, although this study was performed on a low number of tissues [84].

The possibility that ovarian tumor microenvironment may be a source of splicing isoforms markers raise interesting questions regarding the studies conducted on whole tumors. First, some of the RNA transcripts detected may actually come from the microenvironment cells. For example, fibulin and fibronectin are two ECM components known to be produced and secreted by stromal cells. Pinpointing the cell type that produced those splicing deregulated secreted proteins will certainly help to rationalize the complex autocrine and paracrine pathways implicated in the cell to cell communication that take place into and surrounding the ovarian tumor. Second, AS is a highly tissue-specific process, some of the splicing pattern changes might be the reflection of the different proportion of stromal and epithelial cells of ovarian tumors. Theoretically, those effects would be minimal when ovarian tumors of equivalent epithelial content (typically 50-75%) are compared but maximal when normal
ovaries (1% epithelial cells) are used as normal reference. As a consequence, prognosis marker derived from cancer samples comparison should yield more reliable splicing markers than diagnosis marker normalized with normal ovaries.

5.3. Alternative splicing associated protein markers

Interestingly, a number of RNA splicing isoforms markers might be amenable to detection at the protein level using isoform-specific antibodies. Ultimately, these could serve as diagnostic or prognostic tool to either directly detect the presence of cancer cells or indirectly the protein in patient’s fluid. Indeed, the product of the genes encoding fibronectin 1, fibulin, osteopontin, galectin 9, platelet derived growth factor A, extracellular sulfatase 2 and slit homolog 2 are all secreted in the extracellular matrix. Even some cytosolic proteins such as utrophin and serine hydroxymethyltransferase 1 were found in patient’s serum [89]. Others are cell surface protein (amyloid beta A4 protein, stromal interaction molecule 1, CD97, peptidyl-glycine alpha-amidating monooxygenase and chemokine-like factor) harboring an ASE that encodes for an extracellular domain. More impressively, the exon encoding the transmembrane domain of betacellulin is preferentially excluded in ovarian tumors [44], leading to a secreted version of the protein [90]. Thus in every cases, isoform-specific antibodies could be theoretically raised against the cancer associated isoform to ultimately serve as diagnostic/prognostic tool to either detect cancer cells or detect the protein in patient’s fluid.

Inversely, the splicing isoforms of the cell surface receptor Fas and CD44 were mostly studied at the protein level by either immunohistochemistry (IHC) or ELISA. Fas linked extracellular apoptotic signals that converge to the programmed cell death pathway through caspase 8 and 10. Differential usage of exon 6, which encodes the single pass transmembrane domain, results in a soluble version (sFas) and a membrane anchored version (mFas). The level of sFas is increase in ovarian tumor of higher grade compared to low grade [91,92] and correlates with worst prognosis for these patient [91]. Although these studies were performed in small cohort, it elegantly demonstrated that AS can produce isoforms detectable in patient’s serum.

The glycoprotein CD44 is a cell surface receptor that binds diverse extracellular matrix ligands such as hyaluronic acid, fibronectin, osteopontin, collagen and laminin [93]. The binding of low molecular weight hyaluronan polymer promotes the motility and invasion properties of CD44 (review in [93]). It is encoded as a 20 exons gene that exhibit extensive AS of the extracellular domain of exons 6 to 15 (also called variable exons 1 to 10). The major isoform present in normal epithelial [94,95] or stromal [96] ovarian cells is the shorter isoform CD44s lacking all variable exons (CD44s for standard isoform). In contrast, a complex pattern of splicing isoforms were detected in cancer tissues, including most of ovarian tumors by mean of RT-PCR [94,97,98] or by IHC using isoform specific antibodies [95,96,99,100]. One of these splicing isoforms, the inclusion of exon v10, appears to correlate with prognosis and is indicative of improved survival in a multivariate analysis of a 142 patient cohort by IHC [96]. However, these findings contrast the initial study of Schroder who found no exon v10 expression although it relies on a smaller cohort [100]. Intriguingly, inclusion of exon v10 in metastatic tumors was correlated with decrease survival [96]. This ap-
parent discrepancy could be rationalized if the exon v10 inclusion is seen as crucial to maintain proper cell adhesion and avoid cell detachment [101]. It remains to be determined if any of the variable exons of CD44 could serve as biomarker at the RNA level.

6. Concluding remarks

AS dramatically increase the diversity of protein expression in human cells and therefore exponentially increase the number of potential disease markers. However, the complexity in detecting AS and the unclear function of the majority of splice variants greatly reduced the rate of AS based ovarian cancer biomarkers. This trend is likely to change in the next few years with the explosion of whole transcriptome sequencing efforts and the inevitable identification of splice variants as byproducts of next generations’ expression profiles. The real challenge now is to develop techniques allowing the use of splicing markers in the clinic and prepare pathologists to this new wave. Clearly, a compelling argument is needed to drive this drastic change in clinical practice and it will most likely be driven by the success of AS based screens in rationally predicting secreted protein that may serve as non-invasive ovarian cancer markers.

Author details

Jean-Philippe Brosseau\textsuperscript{1,2} and Sherif Abou-Elela\textsuperscript{1,3}

*Address all correspondence to: sherif.abou.elela@usherbrooke.ca

1 Laboratoire de Génomique Fonctionnelle de l’Université de Sherbrooke, Canada

2 Département de Biochimie, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Québec, Canada

3 Département de Microbiologie et d’Infectiologie, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Québec, Canada

References


[92] Midis GP, Shen Y, Owen-Schaub LB. Elevated soluble Fas (sFas) levels in nonhematopoietic human malignancy. Cancer Res 1996 Sep 1;56[17]:3870-3874.


