Rapid Analysis of Distinctive CD44 RNA Splicing Preferences That Characterize Colonic Tumors

Steven Goodison, Kazuhiro Yoshida, Takashi Sugino, Anthony Woodman, Hazel Gorham, John Bolodeoku, Martin Kaufmann, and David Tarin

Nuffield Department of Pathology and Bacteriology, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom (S. G., K. Y., T. S., A. W., H. G., J. B., D. T.), and R&D New Diagnostics, Oncology LP-NIF/TU, Boehringer Mannheim GmbH, Werk Tutzing, Bahnhostrasse 9–15, D-82372 Tutzing, Germany (M. K.)

Abstract

In normal tissues, the steady-state level of CD44 mRNA is low, and the variety of alternatively spliced transcripts produced by this complex gene is limited. Conversely, increased and disorderly expression of this gene has been observed in a number of types of cancer. This study analyzed the order in which the CD44 variant exons are spliced together in gastrointestinal tumor cell lines and in 20 colonic carcinomas and matched normal mucosa.

We used a PCR-based assay to analyze specific exon junctions at the boundary of the standard and variant regions of the CD44 gene transcripts. This revealed characteristic different splicing preferences in colonic tumor and normal tissues. The junction of exon 5 to exon 8 appeared to be the most prevalent in normal mucosa, whereas the presence of junctions between exon 5 and either exon 7, 9, or 11 were increased markedly in tumor samples.

These observations demonstrate that the unusual variety of CD44 transcripts in cancer cells results from the fidelity of alternative splicing mechanisms being compromised and are potentially useful as tumor cell markers in diagnostic assays.

Introduction

The human CD44 gene is composed of at least 20 exons, of which 10 are expressed consistently together to encode for a peptide backbone of about M, 37,000, which may subsequently be glycosylated before being expressed as the "standard isoform" (CD44v) on the cell surface (1, 2). mRNA transcripts can contain the other variably expressed exons of the gene in various combinations with the CD44s exons in a tissue-specific manner, and this is determined by alternative splicing mechanisms. This results in the production of many isoforms, presumably with different functions, from the same gene (3, 4).

Severe disturbances in the pattern of CD44 expression have been observed in a variety of human tumors by both protein and RNA analyses. Markedly increased overexpression levels of CD44 transcripts and proteins have been recognized in many tumors compared to normal cells from the same tissue (5–7). With RT-PCR/hybridization analysis using probes for CD44 exons, distinctive, large molecular weight amplicons are visualized as well as a markedly overall increase in the quantity of CD44 mRNA, specifically in tumor cell-containing samples. In many such samples, a very wide range of sizes of CD44 mRNA species are evident, indicating a severely deranged array of gene transcripts. These characteristic abnormalities may be an indirect result of transcriptional deregulation, which leads to overexpression of the CD44 gene. This may in turn lead to subsequent leakage through the splicing mechanisms in tumor cells, resulting in accumulation of immature or defectively processed mRNA transcript species. Accordingly, we have recently observed the retention of introns in mature mRNA transcripts (8, 9), an event that could result in the production of abnormal or truncated protein products.

Inappropriate expression patterns of the variable exons have been linked both to tumor growth and to metastatic potential (10, 11). When CD44 expression is analyzed in tissue samples by RT-PCR and Southern hybridization with the use of single variant exon-specific probes, it is found that numerous transcripts containing different combinations of variant exons are present in tumor samples but not in the corresponding normal tissues. This finding suggests that some loss of the normal tight regulation of CD44 RNA splicing occurs in tumor cells.

The possible clinical application of these observations has been shown by the detection of abnormal CD44 expression in tissue specimens from tumors of many different histogenetic origins. Furthermore, we have been able to achieve noninvasive detection of malignancy by detailed analysis of CD44 expression in exfoliated cells in body fluids and waste products (12).

These findings stimulated more detailed examination of the splicing pattern of CD44 transcripts to determine whether such a profile could be useful as a diagnostic indicator of carcinoma in a human tissue or fluid sample. This study evaluated the order of exon assembly at the 5' boundary between standard and variant portions of the molecule in gastrointestinal cell lines and compared such patterns in specimens from human colon carcinomas and matched normal tissues.

Materials and Methods

Cell Lines and Tissues. Initially, a number of cell lines were studied (HT-29 and TE-1; Refs. 7 and 12) for the presence of specific CD44 exon junctions. HT-29 is a colonic carcinoma cell line. The TE series of esophageal carcinoma cell lines were provided by Dr. T. Nishihira (Tohoku University, Tohoku, Japan). Cells were routinely cultured in RPMI 1640 (Life Technologies, Inc., Gaithersburg, MD) supplemented with 10% FCS at 37°C in an atmosphere containing 5% CO2.

Fresh colorectal carcinoma tissue samples and corresponding normal colonic mucosa from surgical resection specimens were collected within 20 min of removal and snap frozen and stored in liquid nitrogen until use. The samples were collected as consecutive cases with no previous knowledge of CD44 status. The presence of carcinoma cells in tissues taken from colon resections was routinely confirmed by cryostat sectioning before analysis.

RT-PCR and Southern Hybridization. Total cellular RNA was extracted by the acid guanidium phenol-chloroform method, and mRNA was purified using oligotex dT (QiaGen). cDNA was synthesized with reverse transcriptase followed by amplification by PCR as described previously (13). Five μg of total RNA or 100 ng of polyadenylated, selected RNA were used for the reverse transcriptase reaction with the cDNA Cycle Kit (Invitrogen). The conditions of PCR were as follows: 94°C for 5 min and 83°C for 1 min (at which time Taq polymerase was added) followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min. The sequences of primers used for amplification across the entire variant region were as follows: P1, 5'-(GACA-
efficient polymerase extension is dependent on stable annealing to the target binding and extension of the primer in the absence of a variant exon, because that only six 3' bases of the 24-base primer annealed to exon 5. This prevented exon 5, so that the presence of a specific variant would result in an amplified TAGTGCTCATCAAAGTGGTAG-3'; 5/8, 5'-TGGTATGAGCTGAGGCTGTGGTAG-3'; CTGGATGGTAG-3'; 5/12, 5'-TGGTATGAGCTGAGGCTGTGGTAG-3'; and 5/14, 5'-TCT GCTGTC-3'; and ADI, 5'-GGTGCTG...nts...exon 11: T6, 5'-TCCAGGCAACTC...and P4, complementary to exons 3 and 17, respectively. Hybridization probes for individual variant exons were generated by PCR using exon-specific primers to amplify sequences from a human CD44 genomic clone template (C2311). UTR, untranslated region.

CATATTGCTTCAATGCTCAGC-3', and P4, 5'-GATGCCAAGATGATCAGTCACAGAC-3', and ADI, 5'-GGTGCTG...served when the presence of exon 7 was confirmed by a Pl-E2 amplification, verifying the presence of a complex array of misprocessed CD44 transcripts. Over-exposure of the autoradiograph allowed the detection of weak signals and the identification of the less abundant, larger-transcript species. The tumor samples regularly contained more of the higher molecular weight transcripts than matched normal tissue samples. By the use of probes for both exons 12 and 8, hybridization signals were obtained in 16 of 20 (80%) tumor samples. With the exon 12 probe, signals were detected weakly in only 3 of 20 (15%) normal mucosae, whereas exon 8 expression was seen in 14 of 20 (70%) matched normal tissue samples (Fig. 3). Expression of exon 7 was detected in 54% of tumor tissue samples but in only 15% of matched normal mucosae.

Optimization of CD44 Exon Junction Analysis Using Esophageal and Gastrointestinal Tumor Cell Lines. A schematic representation of the PCR-based approach is shown in Fig. 2. Initial studies were performed on a number of cell lines to optimize the “exon-link” assay conditions for experimental analysis of the identity and the pattern of assembly of variant CD44 transcripts in tissue samples. In three esophageal cell lines (TE series) and the colonic carcinoma line (HT-29), every possible 5/X junction was found to be present in CD44 mRNA. A primer (E2) complementary to 3’ sequences of exon 7 was used with the anchor primer P1 in a control reaction, and this performed several functions. A P1–5/7 amplicon was only seen when the presence of exon 7 was confirmed by a P1-E2 amplification, verifying the specificity of the 5/7 junction primer, and therefore of all junction primer design. Also, the fact that the P1-E2 product was always of the expected molecular weight (420 bp) confirmed that CD44 exon 6 that is expressed in mice was not included in human tissue transcripts. These findings also showed that a putative exon 5 “cryptic” splice site (16) was not used in the tissues thus far tested. A plasmid construct (pBS1S) derived from peripheral blood lymphocytes was transfected into human cell lines, and the entire variant exons by PCR using primers P1 and P4, complementary to exons 3 and 17, respectively. Hybridization probes for individual variant exons were generated by PCR using exon-specific primers to amplify sequences from a human CD44 genomic clone template (C2311). UTR, untranslated region.

CD44 Exon-specific RT-PCR/Southern Hybridization Analysis. In accordance with our previous studies (14, 15), PCR amplification resulting from the use of primer pair P1-P4 on tumor cell lines and tissue cDNA produced a product of 482 bp in all samples representing the ubiquitously expressed CD44s transcript (exons 1–5 and 16–20). In tumor samples, this product was accompanied by several amplicons of higher molecular weight, and hybridization with exon-specific probes confirmed that the corresponding abundant transcripts contained combinations of several variant exons. The expression of total CD44 transcripts in all samples was assessed by hybridization of the blotted PCR product with a probe complementary to exons 4 and 5 (standard exon probe), shown in Fig. 3. On hybridization with the standard probe and certain variant exon probes, many tumor samples gave such strong signals that they appear as a dense smear on the resulting autoradiograph, suggesting the presence of a complex array of misprocessed CD44 transcripts. Over-exposure of the autoradiograph allowed the detection of weak signals and the identification of the less abundant, larger-transcript species. The tumor samples regularly contained more of the higher molecular weight transcripts than matched normal tissue samples. By the use of probes for both exons 12 and 8, hybridization signals were obtained in 16 of 20 (80%) tumor samples. With the exon 12 probe, signals were detected weakly in only 3 of 20 (15%) normal mucosae, whereas exon 8 expression was seen in 14 of 20 (70%) matched normal tissue samples (Fig. 3). Expression of exon 7 was detected in 54% of tumor tissue samples but in only 15% of matched normal mucosae.

Optimization of CD44 Exon Junction Analysis Using Esophageal and Gastrointestinal Tumor Cell Lines. A schematic representation of the PCR-based approach is shown in Fig. 2. Initial studies were performed on a number of cell lines to optimize the “exon-link” assay conditions for experimental analysis of the identity and the pattern of assembly of variant CD44 transcripts in tissue samples. In three esophageal cell lines (TE series) and the colonic carcinoma line (HT-29), every possible 5/X junction was found to be present in CD44 mRNA. A primer (E2) complementary to 3’ sequences of exon 7 was used with the anchor primer P1 in a control reaction, and this performed several functions. A P1–5/7 amplicon was only seen when the presence of exon 7 was confirmed by a P1-E2 amplification, verifying the specificity of the 5/7 junction primer, and therefore of all junction primer design. Also, the fact that the P1-E2 product was always of the expected molecular weight (420 bp) confirmed that CD44 exon 6 that is expressed in mice was not included in human tissue transcripts. These findings also showed that a putative exon 5 “cryptic” splice site (16) was not used in the tissues thus far tested. A plasmid construct (pBS1S) derived from peripheral blood lymphocytes was transfected into human cell lines, and the entire variant exons by PCR using primers P1 and P4, complementary to exons 3 and 17, respectively. Hybridization probes for individual variant exons were generated by PCR using exon-specific primers to amplify sequences from a human CD44 genomic clone template (C2311). UTR, untranslated region.

CD44 Exon-specific RT-PCR/Southern Hybridization Analysis. In accordance with our previous studies (14, 15), PCR amplification resulting from the use of primer pair P1-P4 on tumor cell lines and tissue cDNA produced a product of 482 bp in all samples representing the ubiquitously expressed CD44s transcript (exons 1–5 and 16–20). In tumor samples, this product was accompanied by several amplicons of higher molecular weight, and hybridization with exon-specific probes confirmed that the corresponding abundant transcripts contained combinations of several variant exons. The expression of total CD44 transcripts in all samples was assessed by hybridization of the blotted PCR product with a probe complementary to exons 4 and 5 (standard exon probe), shown in Fig. 3. On hybridization with the standard probe and certain variant exon probes, many tumor samples gave such strong signals that they appear as a dense smear on the resulting autoradiograph, suggesting the presence of a complex array of misprocessed CD44 transcripts. Over-exposure of the autoradiograph allowed the detection of weak signals and the identification of the less abundant, larger-transcript species. The tumor samples regularly contained more of the higher molecular weight transcripts than matched normal tissue samples. By the use of probes for both exons 12 and 8, hybridization signals were obtained in 16 of 20 (80%) tumor samples. With the exon 12 probe, signals were detected weakly in only 3 of 20 (15%) normal mucosae, whereas exon 8 expression was seen in 14 of 20 (70%) matched normal tissue samples (Fig. 3). Expression of exon 7 was detected in 54% of tumor tissue samples but in only 15% of matched normal mucosae.

Optimization of CD44 Exon Junction Analysis Using Esophageal and Gastrointestinal Tumor Cell Lines. A schematic representation of the PCR-based approach is shown in Fig. 2. Initial studies were performed on a number of cell lines to optimize the “exon-link” assay conditions for experimental analysis of the identity and the pattern of assembly of variant CD44 transcripts in tissue samples. In three esophageal cell lines (TE series) and the colonic carcinoma line (HT-29), every possible 5/X junction was found to be present in CD44 mRNA. A primer (E2) complementary to 3’ sequences of exon 7 was used with the anchor primer P1 in a control reaction, and this performed several functions. A P1–5/7 amplicon was only seen when the presence of exon 7 was confirmed by a P1-E2 amplification, verifying the specificity of the 5/7 junction primer, and therefore of all junction primer design. Also, the fact that the P1-E2 product was always of the expected molecular weight (420 bp) confirmed that CD44 exon 6 that is expressed in mice was not included in human tissue transcripts. These findings also showed that a putative exon 5 “cryptic” splice site (16) was not used in the tissues thus far tested. A plasmid construct (pBS1S) derived from peripheral blood lymphocytes was transfected into human cell lines, and the entire variant exons by PCR using primers P1 and P4, complementary to exons 3 and 17, respectively. Hybridization probes for individual variant exons were generated by PCR using exon-specific primers to amplify sequences from a human CD44 genomic clone template (C2311). UTR, untranslated region.
phocyte mRNA containing only CD44s exons and therefore only exon junction 5/16 was used as a negative control template in all experiments to confirm the specificity of 5/X primer annealing. A P1-P4-positive PCR control reaction was included with each sample, and this confirmed the presence of CD44 cDNA (Fig. 4).

**Exon-Link Analysis of Colonic Tumor and Normal Mucosa.** Optimization and standardization of the assay was followed by evaluation of CD44 variant transcript assembly in fresh human colonic tissue samples from surgical resection specimens. An example of results obtained with a selection of P1–5/X primer pairs on matched tumor and normal tissues from one resected colon are shown in Fig. 4. The production of amplicons of the expected size (342 bp) confirms the presence within this tumor of a variety of transcripts in which exons 7, 8, and 9 abut on the standard exon 5, whereas only the junction of exon 5 with exon 8 is detected in the normal tissue sample.

The results of the full exon-link analysis of 20 tumor and 20 matched normal mucosal samples are presented in Table 1, and a number of observations emerge from the data. Colonic tumor and normal tissues both displayed 80% prevalence of junction 5/8 (16 of 20 cases), and this correlates well with the RT-PCR/Southern hybridization results on the same samples (Fig. 3). Together, these data show that junction 5/8 is the most common in normal mucosa and may therefore be the correct, or preferred splicing option in colonic mucosa. Junctions 5/7 and 5/11 are increased markedly in colonic tumor tissues (both 17 of 20 (85%)) compared to their normal counterparts. Both are relatively uncommon in normal mucosa, with scores of 5 of 20 (25%) and 6 of 20 (30%), respectively. Conversely, junction 5/9 was present in 60% of tumors and 35% of normal tissues, making it less suitable as a distinguishing marker. Junction 5/10 appeared to be extremely rare in neoplastic or normal colonic tissues, although it was present in cell lines. Junctions between exon 5 and those toward the 3'-end of the variant region were less prevalent in both normal and neoplastic tissue. The data show that the probability of an exon being spliced against the standard exon 5 diminishes as the exon resides at a more distal site in the pre-mRNA molecule. Junctions between exon 12 or 13 and exon 5 were present in tumor samples at between 15 and 30% prevalence and in normal samples at 10–15%. However, junction 5/14 was found in 40–50% in tumor and normal samples.

**Discussion**

The data obtained from this investigation demonstrate a general disorder involving assembly of CD44 gene transcripts in tumor cells and tissues. This work was designed to identify whether there is a tumor-defining splicing pattern that could be diagnostically useful and provide insights into the mechanisms involved in the abnormal expression of CD44 in the neoplastic process. In fact, it was found that normal tissues have preferential splicing patterns but that in tumor cells, the regulation of transcript assembly is disorganized, resulting in the generation of a wide variety of unusual, alternatively spliced mRNA species.

In normal colonic mucosa, the standard exon 5 was characteristically adjacent to variant exon 8, but in tumor tissue, transcripts in which exon 5 abuts on variant exon 7 were common. Therefore, because exon 7 is usually removed from transcripts in normal mucosa, its presence is a sign of disturbed CD44 mRNA assembly indicative of neoplasia.

From a diagnostic standpoint, RT-PCR/Southern hybridization studies have previously shown that there is both increased transcription and overexpression of the variant exons of this gene in tumor cells. However, this method is technically difficult and laborious for
CD44 ALTERNATIVE SPlicing IN COlONIC TUMOR TISSUE

Table 1  Results of exon-link assay on human cell lines and human colonic carcinoma tissues and corresponding normal colonic mucosa.¹

<table>
<thead>
<tr>
<th>Exon junction</th>
<th>5/7</th>
<th>5/8</th>
<th>5/9</th>
<th>5/10</th>
<th>5/11</th>
<th>5/12</th>
<th>5/13</th>
<th>5/14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell line²</td>
<td>TE-1</td>
<td>HT-29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Normal</td>
<td>17 of 20 (85%)</td>
<td>16 of 20 (80%)</td>
<td>12 of 20 (60%)</td>
<td>4 of 20 (20%)</td>
<td>17 of 20 (85%)</td>
<td>3 of 20 (15%)</td>
<td>6 of 20 (30%)</td>
<td>8 of 20 (40%)</td>
</tr>
</tbody>
</table>

¹ Each tissue sample was tested for the presence of each of the possible junctions between the standard exon 5 and variant exons 7–14 in CD44 transcripts.
² TE-1, human esophageal cell line; HT-29, human colon carcinoma cell line.

routine clinical use. This new assay has the advantage of being simpler, in that blotting and hybridization are unnecessary, without compromising the specificity of detection of CD44 mRNA species. The comprehensive analysis of exon junctions at the boundary between the standard and variant regions described above has demonstrated that the combinations 5/7 and 5/11 are the most prevalent in tumor tissues and therefore of the most interest. These junctions were seen in 85% of tumors, and this value compares favorably with the best tumor marker data obtained with the more complicated methods used previously for CD44 expression analysis (9, 15).

To investigate the possibility that splicing derangement in tumor tissue could be a general phenomenon, we assessed the gross splicing pattern of three other alternatively spliced genes in the same cDNA samples. Transcripts of the genes fibronectin, tenascin, and CaM kinase II were monitored by RT-PCR using amplification across alternatively spliced regions (17–19), but such analysis revealed no gross derangement of splicing pattern of these genes between the tumor and normal matched tissues tested (data not shown). In contrast, the pattern obtained with the CD44-specific P1-P4 primers (Fig. 3, standard probe) is markedly different between tumor and normal colonic mucosa.

As well as the identification of a specific exon junction for use in tumor detection, the data obtained in this study also allow a number of observations to be made regarding CD44 expression. The observation that all possible junctions tested were present in the carcinoma cell lines without any typical patterns demonstrates that the widespread disorganization of CD44 expression and alternative splicing in such source material is not truly representative of tumors in vivo. The cell lines were useful as positive control material, especially in the optimization of the assay and the verification of all primers. However, for the clinical assessment of a new observation, it is ultimately necessary to examine human tissue directly.

The sensitivity of PCR is such that the presence of relatively few molecules containing the junction under investigation would be detected. Therefore, the absence of a signal with a given primer combination in such an assay indicates that the corresponding exon junction is extremely rare. The presence of an exon-5 to exon-7 junction in 10–15% of normal colonic tissues may represent a “background” of misprocessed CD44 transcripts in some individuals. Given that no biological process is 100% efficient, there may be considerable variation in the degree of splicing efficiency between individuals in a population.

Although splicing is more disorganized in the tumor tissue samples, the data suggest certain patterns and preferences of splicing in the nuclei of colonic mucosal cells. For instance, exon 8 seems to be the most common variant exon abutting exon 5, whereas exon 10 appears to be spliced out of the pre-mRNA efficiently in both tumor and normal tissues. It may be that this exon is the most accessible and therefore the first splicing site engaged. The data also suggest that other exons appear to be spliced together as pairs or as cassettes. For example, it has previously been documented that exons 11 and 12 are equally discriminatory for tumor detection in colonic mucosa by Southern hybridization analysis (Ref. 9, Fig. 3). The exon-link data showing the markedly increased presence of junction 5/11 compared to junction 5/12 implies that they are spliced as a unit together in ascending order.

The data obtained in this work suggest some possible new avenues for tumor diagnostic applications using the abnormalities in CD44 gene expression. These include immunoassays for exon-spanning epitopes encoded by transcripts prevalent in tumor tissues. Additionally, a test for tumor-related exon junctions could be incorporated in multiplex assays for other known disorders of CD44 expression, such as intron retention (15). Analysis of the ratios of unusual exon junctions to those seen typically in a given normal tissue could be accomplished by automated, PCR-based analytical methods now becoming available. This could conceivably provide a rapid, clinically relevant digital result.

Acknowledgments

We thank Linda Summerville for preparation of the manuscript and Helene Mellor, Susannah Crowley, and Heather Dorricott for their technical assistance. We would also like to thank Dr. Heinz Bodenmuller for helpful scientific discussions.

References

Rapid Analysis of Distinctive CD44 RNA Splicing Preferences That Characterize Colonic Tumors

Steven Goodison, Kazuhiro Yoshida, Takashi Sugino, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/57/15/3140

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/57/15/3140. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.