In-silico Analysis of Kallikrein Gene Expression in Pancreatic and Colon Cancers

GEORGE M. YOUSEF1,2, CARLA A. BORGOÑO1,2, CYNTHIA POPALIS2, GEORGE M. YACOUB3, MARY-ELLEN POLYMERIS2, ANTONINUS SOOSAPILLAI1 and ELEFTHERIOS P. DIAMANDIS1,2

1Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario; 2Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada; 3University of Virginia School of Medicine, Roanoke-Salem Internal Medicine Program, Roanoke, VA 24033, U.S.A.

Abstract. Human kallikreins are a cluster of 15 serine protease genes located in the chromosomal band 19q13.4, a non-randomly rearranged region in many solid tumors, including pancreatic cancer. We utilized the SAGE and EST databases of the Cancer Genome Anatomy Project to perform in-silico analysis of kallikrein gene expression in normal and cancerous pancreatic and colon tissues and cell lines using virtual Northern blotting (VNB), digital differential display (DDD) and X-profiler. At least two kallikreins, KLK6 and KLK10, are significantly up-regulated in pancreatic cancer. We probed 2 normal and 6 pancreatic cancer SAGE libraries with gene-specific tags for each of these kallikreins. KLK6 was found to be expressed in 5/6 cancer libraries and showed the most marked (5-fold) increase in average expression levels in cancer vs. normal. These data were verified by screening the EST databases, where all mRNA clones isolated were from cancerous libraries, with no clones detected in normal pancreatic tissues or cell lines. X-profiler comparison of two pools of normal and cancerous pancreatic libraries further verified the significant increase of KLK6 expression levels in pancreatic cancer. DDD data showed a 13-fold increase in KLK10 expression in pancreatic cancer. Three kallikrein genes, KLK6, 8 and 10 are overexpressed in colon cancer compared to normal colon, while one kallikrein, KLK1, is down-regulated. While no expression of KLK6 was detected in normal colon, KLK6-specific tags were detectable in 2 cancer libraries. Similar results were obtained by EST screening: no KLK6 clones were detected in any of the 28 normal libraries examined, while 10 KLK6 EST clones were found in colon adenocarcinoma. KLK10 was not detectable in normal colon. Gene-specific tags were, however, detectable with high density in colon cancer and 7 EST clones were found to be expressed in colon Adenocarcinoma.

Pancreatic cancer is the fifth leading cause of cancer-related deaths. It is usually diagnosed when the tumor has already spread locally or metastasized (1). The overall survival rate of patients with pancreatic cancer is only 3% (2, 3). However, patients with small tumors and those with a successful surgery have a higher survival rate of ~40% (4). Thus, early detection is crucial.

Significant prognostic factors in pancreatic cancer include tumor size, grade, stage and resection margin status (5). In addition, a number of circulating tumor markers are available for pancreatic cancer diagnosis and prognosis, including CA 19-9, CA 242, MUC 1 and Span-1. However, none possess the desirable sensitivity or specificity (5, 6).

Colon cancer is the third most common malignancy and the second leading cause of cancer-related deaths in the United States (7). The lifetime risk for colon cancer is 5-6% and is influenced by the heterogeneous etiology of the disease, involving genetic and environmental factors (7, 8). About 20% of colon cancer cases are attributed to two main hereditary syndromes, namely familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC), while the rest are sporadic (8).

Since patients with early stage colon cancer have an estimated 5-year survival rate of 91%, compared to only 6% for those with later stage disease, early detection remains the most important factor in improving long-term survival. The use of fecal blood testing and colonoscopy in the
screening/diagnosis of colon cancer has increased the overall 5-year survival rate from 41% in the 1950s to 54% in the 1980s (9). Current serum-based tumor markers, such as carcinoembryonic antigen (CEA), CA 19-9, CA 242 and tissue polypeptide antigen (TPA), have a limited role in screening and early diagnosis due to a lack of specificity and sensitivity (10).

For many decades, the Dukes’ classification and the TNM staging system have been the gold standards for predicting outcome and implementing therapeutic strategies in the management of colon cancer patients (11). However, further sub-staging of colon cancer is of importance for prognosis and treatment, particularly for patients with stage II disease, in which 40-50% have aggressive tumors and might benefit from adjuvant chemotherapy (8). Among the numerous serum and cell/tissue-based prognostic/predictive tumor markers recently identified, including CEA, oncogenes (ras, her-2/neu), tumor suppressor genes (p53, p27), growth factors (vascular endothelial growth factor), proteases (matrix metalloproteinases, urokinase-type plasminogen activator) and cell adhesion proteins (CD44, E-cadherin) (12), the most widely applied has been CEA. Its measurement has been recommended for staging, prognosis, detecting disease recurrence, monitoring response to therapy and screening for hepatic metastases (13).

Kallikreins are a group of 15 serine proteases whose genes are clustered together in an area of approximately 300 kb on chromosome 19q13.4 (14, 15), a hotspot region for cancer (16). The relationship of kallikreins with malignancy is well-established (14, 15, 17, 18). Human kallikrein 3, also known as prostate specific antigen (PSA, hK3), a member of this family, is routinely used as a tumor marker for prostate cancer (19). Accumulating evidence in the past few years has revealed the potential role of additional kallikreins in cancer diagnosis and prognosis (20-29).

Tissue kallikrein activity in the gastrointestinal tract was first identified by Werle in 1960 (30) and has been studied in rats, cats and humans (31). The hK1 protein was immunohistochemically localized in the goblet cells (32). Although the biological function of kallikreins in the colon is not known, it has been suggested that they may serve as processing enzymes for mucopolysaccharides, stimulate ion transport across the epithelium of the gastrointestinal system, or be involved in the regulation of local blood flow (32).

The Cancer Genome Anatomy Project (CGAP) is an international effort implemented by the National Cancer Institute (NCI) to create a catalog of the genes associated with cancer and to develop technological tools to support the analysis of the molecular profiles of cancer cells and their normal counterparts (33). Gene expression data contained in the electronic databases at the CGAP can be used to identify potentially informative marker genes expressed in cancer (34) and to compare cDNA frequencies of genes of interest in normal vs. cancer. The project applies two main approaches: the expressed sequence tag (EST) method (35) and the serial analysis of gene expression (SAGE) approach (36). Various analytical tools have also been developed for data analysis including BLAST, Virtual Northern Analysis (VNA), X-profiler and Digital Differential Display (DDD). These databases have been successfully used in recent years and results obtained from in-silico analysis were experimentally verified in most cases (37-39).

In our efforts to analyze differential kallikrein gene expression in various malignancies, we previously reported an in-silico analysis of kallikreins in ovarian (40) and breast cancers (our unpublished data). In this study, we utilized the databases and analysis tools available from the Cancer Genome Anatomy Project (41) to analyze kallikrein gene expression in normal and cancerous pancreatic and colon tissues. The independent EST and SAGE databases were screened for kallikrein expression patterns. Our results indicate that two kallikreins, KLK6 and KLK10, are up-regulated in pancreatic and colon cancer. Other kallikreins are also differentially regulated, though to a lesser extent, in these two malignancies.

Materials and Methods

Serial analysis of gene expression (SAGE) and Virtual Northern blotting (VNB). All publicly available SAGE data up to and including December 2002 were used for analysis of kallikrein gene expression. Both NlaIII and Sau3A tags from SAGEmap (http://www.ncbi.nlm.nih.gov/SAGE/) were mapped to UniGene clusters (http://www.ncbi.nlm.nih.gov/UniGene/), available through the NCBI website (http://www.ncbi.nlm.nih.gov/). Each UniGene group consists of all GeneBank sequences representing the same human gene. Hereafter, each such group will be referred to as a "gene". Tags mapping to more than one gene were excluded.

The mRNA sequences of the 15 human kallikrein genes, obtained from the Human Genome Project, were used to identify unique sequence tags of UniGene clusters for each kallikrein (for GeneBank accession numbers, see references (14, 42, 43)). These sequence tags were then used to determine the levels of expression of different kallikreins in 2 normal and 6 pancreatic cancer libraries and 2 normal and 6 colon adenocarcinoma libraries. A list of the tags used for analysis is provided in Table I and detailed information about these libraries are available from the website of the Cancer Genome Anatomy Project (CGAP) (http://www.ncbi.nlm.nih.gov/ncicgap/). Analyses were performed by comparing the proportion of libraries of each type (cancer vs. normal) that show expression of each tag in addition to the average expression densities within each library. If more than one tag of the same gene appeared in the same library, we only included the one with the highest expression density (maximum tpm) and the other tag was excluded to avoid inaccurate estimation of expression. Expression levels are displayed as blots with different densities and corrected as tpm to facilitate comparison (Figure 1).

Expressed sequence tag (EST) analysis. The full-length mRNA sequence of each kallikrein was compared against the human EST databases of the NCBI. At the time of the study, these databases...
included 8 normal and 11 cancerous pancreatic libraries and 28 normal and 53 colon adenocarcinoma libraries. Expression was calculated for each kallikrein as the number of positive libraries out of the total in each tissue type, in addition to the total number of clones detected in each type.

**X-profiler and digital differential display (DDD) analysis.** A comparison of normal and cancerous pancreatic and colon libraries available in the SAGE databases was performed using X-profiler analysis. As expression levels of various kallikreins might be different from one cell line to another and in different types of pancreatic cancer, we compared the two pools of all normal and cancerous libraries available in the databases. The X-profiler cut-off value was set at 2-fold difference.

The DDD search engine (33) was used to compare EST expression in normal and cancer libraries. The databases at the time of analysis included 8 normal and 11 pancreatic cancer libraries. Libraries with less than 25 clones were excluded from this study.

### Results

**Preliminary evaluation of in-silico analysis.** We first verified the reliability of in-silico analysis by comparing normal kallikrein tissue expression patterns obtained by in-silico analysis with previously published experimental results (42, 44-49). SAGE and EST results were in general agreement with PCR and Northern blotting data. For example, both in-silico and experimental data indicate that KLK2-4, KLK5, and KLK7 are highly expressed in the prostate, while KLK5 shows high expression in normal breast tissue and skin (50, 51) (data not shown). In addition, we found that those kallikreins exhibiting differential expression in pancreatic or colon cancer have unchanged levels of expression in other malignant states, further verifying the accuracy of our results.

**Kallikrein expression in pancreatic cancer**

**SAGE and VNB for kallikrein expression in normal and cancerous pancreatic tissues.** Human kallikrein gene 6 (KLK6, zyme/protease M/neurosin) displayed the most significant differential expression in pancreatic cancer compared to normal pancreatic tissues. As shown in Table II, probing all pancreatic libraries from different sources (cancerous and normal from tissues and cell lines) with KLK6-specific mRNA tags revealed that while KLK6 was detectable in only 1 of 2 normal pancreatic libraries with relatively low density (average expression of 31 tpm), mRNA tags were detectable in 5 out of 6 cancer libraries. The expression density was 5-fold higher in cancer (average 155 tpm).

Similarly, KLK10-specific tags were detectable with an average density of 32 tpm in 4 out of 6 pancreatic cancer libraries, but not measurable in any of the 2 normal pancreatic libraries analyzed (Table II). While KLK7 and KLK11 were not found in normal pancreas, their expression was observed in 2 and 3 pancreatic cancer libraries, respectively. The average expression density was 28 tpm for KLK7 and 35 tpm for KLK11.

Other kallikreins were also found to be expressed at slightly lower levels in the pancreas but with no significant difference in expression between normal and cancerous tissues. For example, KLK1 was expressed at comparable levels in both normal and cancerous tissues (Table II). KLK8 was expressed with an average density of 31 tpm in normal pancreatic tissues compared to 28 tpm in malignancy. Similar levels were found for KLK14 (data not shown for the latter two).

**EST analysis of differential kallikrein gene expression in pancreatic cancer.** Our EST library screening results (Table III) provided further verification of SAGE expression and VNB data. KLK1 expression levels were comparable in normal and cancer (three clones were detected in each). Sixteen KLK10 EST clones were detected in 3 out of 11 pancreatic cancer libraries analyzed. Results for KLK6 were not as clear, since only one clone was detected from pancreatic cancer tissues with no clones detected in normal. This low expression might be due to the presence of pancreatic-specific splice variants or a generally low level of KLK6 expression in the pancreas, rendering this gene detectable only in the more quantitative SAGE databases (see below). Results for KLK7 and KLK11 were not informative. Interestingly, KLK7 was also expressed in fetal pancreatic tissues. No other kallikreins were differentially expressed in pancreatic cancer using EST database analysis. It should be realized, however, that quantitative EST figures are only approximate and cannot be relied upon for quantification due to the fact that some EST libraries are normalized.

**X-profiler and DDD analysis of kallikrein gene expression.** Table IV shows the analysis of kallikrein gene expression in normal and cancerous pancreatic tissues utilizing X-profiler and DDD analysis tools. A pool of 2 normal pancreatic tissues and cell lines was compared to another pool of 6 pancreatic cancer tissues and cell lines (all non-normalized) by X-profiler analysis. Our results clearly indicate significant differences in the expression levels of KLK6 in pancreatic cancer compared to normal tissues. An expression factor of 0.806 was detected. When two pools of 8 normal and 10 cancerous EST libraries were compared in the DDD analysis, a 13-fold increase in KLK10 expression was observed in the cancerous libraries. Such a difference was not seen for other kallikreins by DDD, which is acceptable since the DDD engine is based on the EST library databases, many of which are normalized or subtracted.
Table I. Gene-specific SAGE tags used to probe different libraries of the CGAP databases.

<table>
<thead>
<tr>
<th>Kallikrein</th>
<th>Restriction enzyme</th>
<th>SAGEtags</th>
<th>UniGene cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLK1</td>
<td>NlaIII</td>
<td>GGGCTACGTC GTGACAGAGG</td>
<td>Hs.123107</td>
</tr>
<tr>
<td>KLK6</td>
<td>NlaIII</td>
<td>CACTCAATAA GCCGCTCTCTG Sau3A CAAAAAACCA ACAGCCCGGA</td>
<td>Hs.79361</td>
</tr>
<tr>
<td>KLK7</td>
<td>NlaIII</td>
<td>CCCTGTGTGAT</td>
<td>Hs.151254</td>
</tr>
<tr>
<td>KLK8</td>
<td>NlaIII</td>
<td>GTCTTGTCAG Sau3A TCCCTTAATA</td>
<td>Hs.104570</td>
</tr>
<tr>
<td>KLK10</td>
<td>NlaIII</td>
<td>TAAGGGTTTAA Sau3A CAGATGCCCA</td>
<td>Hs.69423</td>
</tr>
<tr>
<td>KLK11</td>
<td>NlaIII</td>
<td>GTGTTGTCGCA Sau3A CAGAGAGCA</td>
<td>Hs.57771</td>
</tr>
</tbody>
</table>

* Only kallikreins that show positive matches are shown in this table.

Table II. In-silico analysis of kallikrein gene expression in normal and cancerous pancreatic and colon tissues and cell lines using SAGEmap.

<table>
<thead>
<tr>
<th>Kallikrein</th>
<th>Library type</th>
<th>Positivity</th>
<th>Average density1</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLK1</td>
<td>Normal pancreas</td>
<td>1/2</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>2/6</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>2/2</td>
<td>191</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>1/6</td>
<td>88</td>
</tr>
<tr>
<td>KLK6</td>
<td>Normal pancreas</td>
<td>1/2</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>5/6</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>2/6</td>
<td>24</td>
</tr>
<tr>
<td>KLK7</td>
<td>Normal pancreas</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>2/6</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>2/6</td>
<td>25</td>
</tr>
<tr>
<td>KLK8</td>
<td>Normal colon</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>1/6</td>
<td>16</td>
</tr>
<tr>
<td>KLK10</td>
<td>Normal pancreas</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>4/6</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>1/6</td>
<td>131</td>
</tr>
<tr>
<td>KLK11</td>
<td>Normal pancreas</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>3/6</td>
<td>35</td>
</tr>
</tbody>
</table>

1 tags per million (tpm)

Table III. In-silico analysis of kallikrein gene expression in normal and malignant pancreatic and colon tissues and cell lines as determined by EST databases.

<table>
<thead>
<tr>
<th>Kallikrein</th>
<th>Library type</th>
<th>Positivity</th>
<th>No. of clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLK1</td>
<td>Normal pancreas</td>
<td>2/8</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>2/11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>2/28</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>2/53</td>
<td>2</td>
</tr>
<tr>
<td>KLK6</td>
<td>Normal pancreas</td>
<td>0/8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>1/11</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>0/28</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>5/53</td>
<td>10</td>
</tr>
<tr>
<td>KLK7</td>
<td>Normal pancreas</td>
<td>1/8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>1/11</td>
<td>2</td>
</tr>
<tr>
<td>KLK10</td>
<td>Normal pancreas</td>
<td>0/8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>3/11</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Normal colon</td>
<td>0/28</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colon cancer</td>
<td>5/53</td>
<td>7</td>
</tr>
<tr>
<td>KLK11</td>
<td>Normal pancreas</td>
<td>0/8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pancreatic cancer</td>
<td>1/11</td>
<td>1</td>
</tr>
</tbody>
</table>

1 tags per million (tpm)
Kallikrein expression in colon cancer. Our results indicate that three kallikrein genes, KLK6, KLK8 and KLK10, are overexpressed in colon cancer compared to normal colon, while one kallikrein, KLK1, is down-regulated. SAGE database screening showed that while no expression of KLK6 was detected in normal colon, KLK6-specific tags were detectable in 2 cancer libraries (average expression 24 tpm) (Table II). These data were further verified by screening the independent EST databases. While no KLK6 clones were detected in any of the 28 normal libraries examined, 10 KLK6 EST clones were found in 5 out of 53 colon adenocarcinoma libraries.

Probing all colon libraries from different sources (normal and adenocarcinoma from tissues and cell lines) with KLK10-specific sequences revealed that KLK10 is not detectable in normal colon as determined by SAGE analysis (Table II) and EST library screening (Table III). Gene-specific tags were, however, detectable with high density (average expression of 131 tpm) in colon cancer and 7 EST clones were found to be expressed in colon adenocarcinoma (Table III). These results were further confirmed by analysis of kallikrein gene expression in normal vs. cancerous colon tissues utilizing the X-profiler tool. A pool of 2 normal tissues was compared against another pool of 6 colon adenocarcinoma libraries (all non-normalized), where an expression factor of 0.613 was detected (Table IV). Such a difference was not seen, however, by DDD, which is acceptable because of the bias effect due to normalization and subtraction.

While KLK7 was not detectable in normal colon, 2 out of 6 cancer libraries showed low expression levels (average 25 tpm). There was also a low expression density of KLK8 in colon cancer, compared to no expression in normal colon. These results were, however, not apparent by EST analysis, presumably due to the generally low expression levels, which were below the detection limit of the EST analysis tool.

KLK1, on the other hand, was found to be down-regulated, at the mRNA level, in colon cancer. VNB analysis showed that KLK1 is expressed in both normal colon libraries examined with an average density of 191 tpm (Table II), compared to positive detection in only 1 out of 6 cancer libraries, with the expression level in cancer being about half (88 tpm). These data were further verified by screening the EST databases where expression levels were found to be approximately doubled in the normal colon (Table III). Moreover, X-profiler analysis showed the same significant difference when comparing KLK1 gene expression between normal and malignant pools of colon libraries, with an expression factor of 0.991 (where a factor of 1.0 represents the highest statistical significance) (Table IV). No significant differences in the expression pattern were found between the distribution of the regular form of KLK1 and the colon-splice variant cloned by Chen et al (52) (data not shown).

KLK11 and KLK15 were found to be expressed at low levels in both normal and cancer colon tissues, with no significant difference in expression levels (data not shown). One adenocarcinoma library showed weak expression of KLK14.

Discussion

The chromosomal band 19q13.4, which harbors the human kallikrein gene locus, is non-randomly rearranged in a variety of solid human tumors, including pancreatic cancer, astrocytomas, ovarian cancer and thyroid tumors (16). Out of the 15 human kallikrein genes, the KLK1 gene was originally named “pancreatic-renal kallikrein” due to its high expression levels in the pancreas (53). KLK1 has been immunolocalized in many pancreatic tissues (54), more specifically in the acinar, β cells and rat transplantable acinar cell carcinoma of the pancreas (32, 55). Its protein product, hK1, can cleave pro-insulin and is implicated in diabetes (32, 56). Harvey et al. (57) have recently reported abundant expression, by Northern blotting, of KLK6-13 in normal pancreatic tissues.

In recent years, many kallikreins have been shown to be differentially regulated in diverse malignancies (15, 20), such as prostate (58), breast (59-61), ovarian (18, 21, 23, 25, 26, 62, 63) and testicular cancer (64, 65). A recent study reported the localization of immunoreactive KLK1 mRNA in pancreatic adenocarcinoma, especially when undifferentiated and pointed to the possible involvement of kallikreins in cancer cell invasiveness (66). Our results indicate that at least two other kallikrein genes (KLK6 and 10) are significantly up-regulated in pancreatic cancer.

Interestingly, although KLK6 mRNA is expressed at a high density in pancreatic cancer according to SAGEmap and X-profiler data analysis, only one EST clone was isolated from a pancreatic cancer cell line. This might be due to the presence of alternatively spliced variant transcripts of KLK6 that are cancer-specific. Anisowicz et al. (67) previously reported the presence of a different splice form of KLK6 mRNA expressed exclusively in the pancreas. Our previous RT-PCR analysis indicated that KLK6 is expressed at very low levels in normal pancreatic tissues (68). The hK6 protein was recently shown, by immunohistochemistry, to be found in the islets of Langerhans of the pancreas (69). Conversely, the hK6 protein was undetectable in the pancreas using a highly sensitive immunofluorometric assay (70).

Previous Northern blot and RT-PCR analysis showed that KLK7 (previously known as the human stratum corneum chymotryptic enzyme, HSCCE) is not measurable in normal pancreas (47, 71). It is thought to play a role in the desquamation of the skin and, more recently, it was shown to be overexpressed in ovarian cancer patients (72).
The KLK10 gene (also known as normal epithelial cell-specific 1, NES1) is expressed in the normal pancreas at the mRNA level (73). Northern blot analysis showed that KLK10 mRNA expressed in the pancreas represents an alternatively spliced, shorter form of the gene. More recently, the hK10 protein was also identified in pancreatic tissues by immunohistochemistry (74). As well, KLK10 and hK10 are differentially expressed in other malignancies including breast, testicular and ovarian cancer (63, 75-77). Our results are consistent with a recently published DDD analysis of genes differentially expressed in a number of solid tumors and listed KLK10 to exhibit 28-fold higher expression in pancreatic cancer (78).

No expression of the KLK11 gene was found, at the mRNA level, in normal pancreatic tissues (44). An alternatively spliced form of the gene was reported in prostate cancer cell lines (79).

Our data also indicates overexpression of KLK6, KLK8 and KLK10 and down-regulation of KLK1 in colon cancer. These results were verified by independent databases and with different analytical tools. Previous RT-PCR studies have identified both KLK6 (73) and KLK10 (73) mRNA in normal colon. These results are consistent with our studies documenting the expression of both hK6 and hK10 proteins in normal colon tissue extracts using ELISA methodologies (70, 80). The hK8 protein, however, was not detected in normal colon by our recently developed ELISA (81). Furthermore, by immunohistochemical analysis, we demonstrated subnuclear cytoplasmic hK6 and hK10 immunostaining in the glandular epithelium of the colon (69, 74). KLK1 mRNA and the hK1 protein were localized in glandular epithelial cells (goblet cells) in colon by in situ hybridization and immunohistochemistry, respectively (32, 82).

Although the underlying biological mechanism of a possible kallikrein involvement in the progression of colon cancer is currently unknown, it is plausible that their effects may be associated with steroid hormones. Epidemiological and experimental evidence indicates that dietary factors, such as phyto-oestrogens, may be protective against colon cancer (83). Such phyto-oestrogens are either inherently estrogenic or converted to estrogenic compounds and may influence sex-hormone production, metabolism and biological activity, intracellular enzymes, protein synthesis, growth factor action, malignant cell proliferation, differentiation, cell adhesion and angiogenesis (83). Since, previous studies have shown that KLK6 and KLK10 are estrogen-regulated genes and that KLK10 may function as a tumor suppressor (68, 75, 84), it is possible that kallikreins are downstream targets in hormonal pathways that affect colon carcinogenesis.

Serine and matrix metalloproteases have been implicated in colon carcinogenesis, by their ability to degrade extracellular matrix (ECM) proteins, thereby facilitating tumor cell proliferation, invasion and metastasis (85). The serine protease, urokinase type plasminogen activator (uPA), promotes cancer invasion and metastasis by converting plasminogen into plasmin which, in turn, helps to degrade the ECM, activate proteases and growth factors (85). Of interest is the fact that hK2 and hK4 are both able to activate the single-chain form of uPA (pro-uPA) in vitro (86, 87). Thus, since KLK6, 8 and 10 encode serine proteases overexpressed in colon cancer, it is not unreasonable to speculate that they may be involved in the promotion of metastasis, directly (by degradation of the basement membrane and ECM) or indirectly (by activation of pro-uPA).

The parallel overexpression of many kallikreins in the same cancer may point out to the possibility of involvement of such proteins in a common pathway that is associated with cancer pathogenesis or progression. The overexpression of the same gene in different, apparently unrelated malignancies is not a phenomenon that is restricted to kallikreins. Many tumor markers, for instance CEA, were found to be elevated in different malignancies. These findings point to a more "biological" rather than "anatomical" classification of cancers.

We used two independent databases (EST and SAGE) to verify our results. We also calculated the proportion of positive libraries from each type in addition to the density of expression. Further confirmation came from comparing our results with the previously reported normal patterns of tissue expression. We did not attempt to quantify kallikrein gene expression from the EST databases because many EST libraries are normalized or subtracted and can produce false results.

The expression of some kallikreins, e.g. KLK6 and KLK11 in the pancreas and KLK8 in the colon, were only detectable in SAGE, but not the EST databases. This observation may be attributed to the fact that these two genes may be expressed in the pancreas at generally low levels, only
Yousef et al: Kallikrein Gene Expression in Pancreatic and Colon Cancer

detectable by the SAGE method, which is able to spot the expression of infrequently expressed or rare genes and to produce much more accurate quantitative data (36).

It should be emphasized, however, that although in-silico analysis is an informative research tool, the results obtained always need experimental verification. Possible sources of bias include sequence errors, presence of specific sequence mutations associated with certain malignancies, unequal representation of different physiological or pathological libraries and the expression of splice variants in certain malignancies.

In conclusion, we provide strong evidence suggesting that at least 3 kallikreins are up-regulated in pancreatic and colon cancers. These data were confirmed from several databases and bioinformatics tools. Further experimental analyses will establish the usefulness of these kallikreins for diagnosis, prognosis and monitoring of patients with pancreatic and colon cancers.

Acknowledgements

This work was supported by a grant to E. P. Diamandis from the Natural Sciences and Engineering Research Council of Canada (NSERC) and IBEX Technologies, through a University-Industry Program.

References


Yousef et al: Kallikrein Gene Expression in Pancreatic and Colon Cancer


Received October 20, 2003
Accepted December 18, 2003