Research Article

TumorBiology

Tumor Biol 2005;26:227–235
DOI: 10.1159/000087377

Received: March 23, 2005
Accepted after revision: April 12, 2005
Published online: August 9, 2005

Identification of New Splice Variants and Differential Expression of the Human Kallikrein 10 Gene, a Candidate Cancer Biomarker

George M. Yousef[a] Nicole M.A. White[a] Iacovos P. Michael[b,c]
Jane Chan-Kyung Cho[c] John Desmond Robb[a] Lisa Kurlender[b,c]
Saba Khan[b] Eleftherios P. Diamandis[b,c]

[a]Discipline of Laboratory Medicine, Memorial University, St. John's, [b]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, and [c]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada

Key Words
Serine proteases · *KLK10* · *NES1* · Breast cancer · Ovarian cancer · Tumor markers · Kallikreins · Serial Analysis of Gene Expression · Cancer Genome Anatomy Project · Splice variants

Abstract
The human kallikrein gene 10 *(KLK10)* is a member of the kallikrein gene family on chromosome 19q13.4. This gene was identified by its downregulation in breast cancer, and preliminary evidence suggests that it may act as a tumor suppressor. A computer-based analysis was performed on EST and SAGE clones from the Cancer Genome Anatomy Project and other databases. Experimental verification of differential expression of *KLK10* in cancer was performed by PCR using gene-specific primers. The mRNA and EST analysis allowed the construction of the longest transcript of the gene and characterization of a 5′ extension of the reported mRNA. In addition, seven new splice variants of *KLK10* were identified. One of these variants, named *KLK10* splice variant 3 *(KLK10-SV3)* which starts with a novel first exon, was experimentally verified. This variant is predicted to encode for the same protein as the 'classical' KLK10 mRNA, since the first exon is untranslated. One variant mRNA partially matches with the sequence of *KLK10*, while the rest of the mRNA matches with a portion of the polycystic kidney disease gene, found on chromosome 15. This variant could not be experimentally verified in either normal or cancerous tissues. There are 39 reported single nucleotide polymorphisms (SNPs) for the gene, in which three result in amino acid substitutions. SAGE analysis shows a clear upregulation of *KLK10* in ovarian, pancreatic, colon, and gastric cancers. The gene is, however, downregulated in breast and prostate cancers. A three-fold decrease in expression levels was noted in actinic keratosis, compared to normal skin from the same patient. The differential regulation of *KLK10* in ovarian and prostate cancers was experimentally verified by RT-PCR analysis. In addition, a significant number of clones were isolated from carcinomas of the head and neck. Fewer clones were found in carcinomas of the skin, brain and prostate. Orthologues were identified in three other species, with the highest degree of homology observed with the mouse and rat orthologues (42% in each). In conclusion new splice variants of the *KLK10* gene were identified. These in silico analyses show a differential expression of the gene in various malignancies and provide the basis for directing experimental efforts to investigate the possible role of the gene as a cancer biomarker.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1010–4283/05/0265–0227$22.00/0

Accessible online at:
www.karger.com/tbi

Dr. E.P. Diamandis
Mount Sinai Hospital, Department of Pathology and Laboratory Medicine
600 University Avenue
Toronto, Ontario M5G 1X5 (Canada)
Tel. +1 416 586 8443, Fax +1 416 586 8628, E-Mail ediamandis@mtsinai.on.ca

Introduction

The human kallikrein gene family is a recently characterized group of serine proteases tandemly located on chromosome 19q13.4 [1]. Members of this family share significant homology at both the gene and protein levels. Recently, all members of this family have been cloned and identified, and it is now established that it consists of 15 genes designated *KLK1–KLK15*. Their corresponding proteins were named hK1–hK15 [2–4].

KLK10, previously known as the normal epithelial cell-specific 1 *(NES1)* gene, was cloned by subtractive hybridization between a radiation-transformed and a normal breast cancer cell line [5]. *KLK10* is expressed in normal mammary epithelial cells but it is dramatically downregulated in breast cancer. *KLK10* spans about 5.5 kb of genomic sequence and is formed of six exons, the first being untranslated, and five intervening introns [6]. The protein product of *KLK10*, hK10, is a predicted secreted serine protease with trypsin-like activity.

Although the biological functions of *KLK10* are unknown, the gene was found to be associated with endocrine-related malignancies such as breast and ovarian cancers. The role of *KLK10* as a candidate tumor suppressor has been reported [7].

In this study, we analyzed all known mRNA clones of *KLK10*, in addition to 130 EST clones and the SAGE databases, to examine the differential expression of the gene in normal and cancerous tissues. We experimentally tested the differential expression of *KLK10* in ovarian and prostate cancers. We also identified seven new splice variants.

Materials and Methods

In silico Analysis

KLK10 ESTs and mRNA clones are obtained from the nonredundant (nr) and EST databases of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) and the University of California at Santa Cruz (UCSC) (http://www.ucsc.edu/).

Information on the *KLK10* gene was obtained from the following databases and web sites:

- The UniGene clusters (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = unigene).
- The OMIM databases (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = OMIM).
- The Swiss-Prot protein knowledgebase (http://us.expasy.org/sprot/).
- The MEROPS databases (http://merops.sanger.ac.uk/).
- The GenCards™ web site (http://genecards.bcgsc.ca/cgi-bin/randomize.pl).
- The Human Gene Nomenclature database (Genew) (http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/searchgenes.pl).
- The CleanEx database of gene expression profiles (http://www.cleanex.isb-sib.ch/).
- The GenAtlas database: human gene database (http://www.dsi.univ-paris5.fr/genatlas/).
- The GeneLynx database; a portal to the human and mouse genomes (http://www.genelynx.org/).
- The Stanford Online Universal Resource for Clones and ESTs (SOURCE) (http://genome-www5.stanford.edu/cgi-bin/SMD/source/sourceSearch).
- The Gene Ontology (GO) database (http://www.geneontology.org/).
- The BLOCKS database; a protein domain database (http://www.blocks.fhcrc.org/).
- The Protein Data Bank; a repository for 3D biological macromolecular structure data (http://www.rcsb.org/pdb/).

Multiple alignments were performed using the 'ClustalW' software package [8], and the BLAST programs of the NCBI, and were manually edited. Splice variants were identified using multiple alignments of sequences obtained from the GenBank, UniGene and the UCSC databases. EST and mRNA sequences were compared to the reference genomic sequence of *KLK10* (GenBank Accession No. AF055481). Alignment viewings were done using the 'boxshade' (www.ch.embnet.org/software/BOX_form.htm l) and 'chroma' (www.lg.ndirect.co.uk/chroma/) programs. Taxonomy information was obtained from the HomoloGene and taxonomy browsers at the NCBI (http://www.ncbi.nlm.nih.gov).

Analysis of *KLK10* gene expression in cancer was performed through the Cancer Genome Anatomy Project (CGAP) databases. EST information was obtained from the GenBank and CGAP web servers. The mRNA sequence of the gene was used to identify unique sequence tags of UniGene clusters and two restriction digestion enzymes (*Nla*III and *Sau*3A) were used as anchoring enzymes. These sequence tags were then used to determine the levels of expression in the SAGE libraries. Detailed information for these libraries is available from the CGAP website (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 in these libraries. If more than one tag of the same gene appears in the same library, we only included the one with the peak level of expression (maximum tags per million, tpm); the other tags were excluded to avoid overestimation of expression.

Reverse Transcriptase Polymerase Chain Reaction

Total RNA was extracted from cell lines and tissues using Trizol reagent (Invitrogen, Carlsbad, Calif., USA) following the manufacturer's instructions. RNA concentration was determined spectrophotometrically. Two micrograms of total RNA were reverse-transcribed into first strand cDNA using the Superscript™ pre-amplification system (Invitrogen). The final volume was 20 μl. Different sets of gene-specific primers were designed for *KLK10* (table 1), and PCR was carried out in a reaction mixture containing 1 μl of cDNA, 10 m*M* Tris-HCl (pH 8.3), 50 m*M* KCl, 1.5 m*M* $MgCl_2$, 200 μ*M* deoxynucleoside triphosphates (dNTPs), 150 ng of primers and 2.5 units of HotStar™ Taq DNA polymerase (Qiagen Inc., Valencia, Calif., USA) on an Eppendorf MasterCycler™. The cycling conditions were 95°C for 15 min to activate the Taq DNA polymerase, followed by 30 cycles of 94°C for 30 s, annealing (temperatures

Table 1. Primers used for PCR amplification of the *KLK10* gene

Experiment	Primer name and sequence[a]		Annealing temp., ° C
Splice variant 3	KLK10-SV3-F KLK10-SV3-R	CTT CCT CCT TCC TCT TCC ACA GTC GTCAGCACCCAACTCTGGTCCAC	67
Translocation variant[b]	KLK10-T-F KLK10-T-R	TAAAGTCATACGACTCCAACTG ACTTTAGTATGTAGTAATAAAGAAC	55.5
Cancer tissue expression[c]	KLK10-TE-F KLK10-TE-R	GGAAACAAGCCACTGTGGGC GAGGATGCCTTGGAGGGTCTC	60

[a] All primers are presented in 5′ → 3′ direction. [b] See figure 1, EST AW999510. [c] See figure 2.

Table 2. Access information of the *KLK10* gene as obtained from different databases

Gene name[a]	Human tissue kallikrein gene 10
Official symbol	*KLK10*
Synonyms	NES1, PRSSL1
GenBank accession	NM_002776 (mRNA) AF055481 (genomic structure)
UniGene Cluster	Hs. 2754641
OMIM	602673
GeneCards ID	GC19M056193
Genew ID	HGNC:6358; KLK10
CleanEx ID	HGNC:6358; KLK10
GeneLynx ID	KLK10; *Homo sapiens:* Hs 6547
GenAtlas ID	KLK10
GO ID	GO:0008236
SOURCE ID	KLK10; *Homo sapiens*
BLOCKS ID	O43240

[a] For full names and web site addresses, please see the Materials and Methods section.

Table 3. Access information of the hK10 protein as obtained from different databases

Official symbol	hK10
Swiss Prot ID	O43240
Merops ID	S01.246
MW[a]	30.1 kDa
Signal peptide	33 aa
Isoelectric point	8.95
Activation peptide	9 aa
Mature protein	234 aa
Cysteine residues[b]	12 aa
Predicted substrate specificity	Trypsin-like

aa = Amino acid.
[a] Excluding post-translational modifications.
[b] Number of residues in the mature enzyme.

shown in table 1) for 30 s, 72°C for 1 min and a final extension step at 72°C for 10 min. Equal amounts of PCR products were electrophoresed on 2% agarose gels and visualized by ethidium bromide staining. All primers for RT-PCR spanned at least two exons to avoid contamination by genomic DNA. To verify the identity of the PCR products, they were cloned into the pCR 2.1-TOPO vector (Invitrogen) according to the manufacturer's instructions. The inserts were sequenced from both directions using vector-specific primers, with an automated DNA sequencer.

Normal and Malignant Tissues

Prostate tissue samples were obtained from seven patients who had undergone radical retropubic prostatectomy for prostatic adenocarcinoma. Fresh prostate tissue samples were obtained from the cancerous and noncancerous parts of the same prostate that had been removed. Histological analysis was performed to ensure that the tissue was either malignant or benign. Normal ovarian cDNA was obtained from the Gene Pool™ cDNA collection (Invitogen). Ovarian tumor tissues were obtained from patients who had undergone surgery for primary ovarian carcinoma.

Results

Analysis of the KLK10 mRNA Sequence and Splice Variants

Access information for the *KLK10* gene and its protein product from different databases are listed in tables 2 and 3. There are five published mRNA sequences for *KLK10* (table 4, fig. 1). These represent two variants of the gene, named *KLK10* splice variant 1 (SV1) and splice variant 2 (SV2). Both variants share the same coding region, with the last four exons (exons 3–6) being identical.

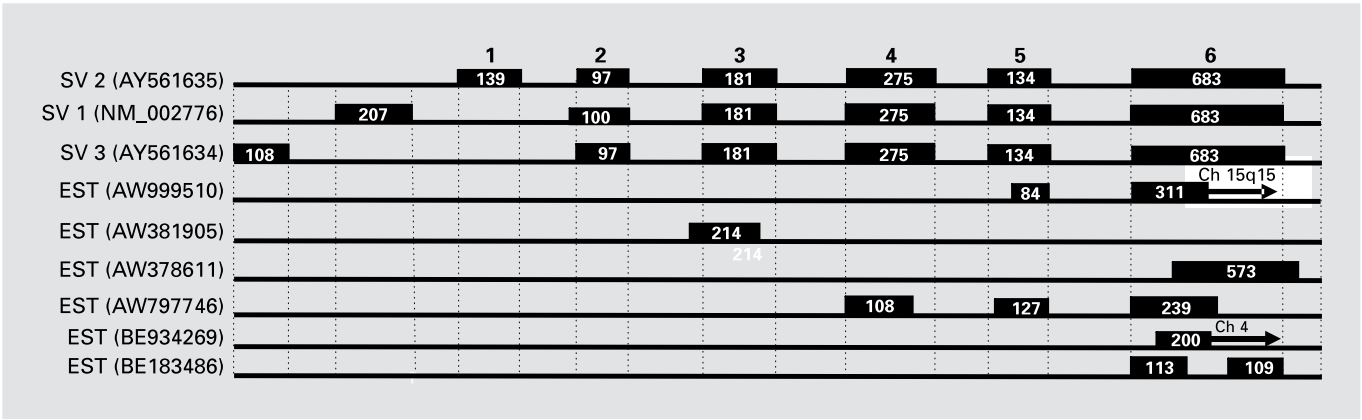

Fig. 1. Diagrammatic representation of *KLK10* splice variants. Solid boxes represent exons and connecting lines are introns. Exon lengths are indicated in base pairs inside boxes. SV = Reported splice variant; EST = potential new variants obtained from EST analysis. For detailed information and full names, see text.

Table 4. Reported *KLK10* mRNA sequences

Variant type	GenBank accession	ID	Length bp	Reference
Splice variant 1	NM_002776	*Homo sapiens* kallikrein 10 (KLK10), transcript variant 1	1,580	NCBI ref. seq.
	BC002710	*Homo sapiens* kallikrein 10	1,542	–
Splice variant 2	AF055481	*Homo sapiens* kallikrein 10	1,442	[6]
	NM_145888	*Homo sapiens* kallikrein 10 (KLK10), transcript variant 2	1,443	NCBI ref. seq.
	AF024605	*Homo sapiens* serine protease-like protease (NES1)	1,454	[5]

They have, however, different first exons. The first exon of KLK10-SV1 is located further upstream (fig. 1). In addition, *KLK10* splice variant 1 has a 3-bp 5′ extension of exon 2. Before our analysis, GenBank accession No. AF024605 was the longest available clone of *KLK10* splice variant 2. However, analyses of 130 *KLK10* ESTs from different databases indicated the presence of five clones with 5′ extension of exon 1 of splice variant 2. The longest clone has a 5′ extension of 66 bp. This 5′-extended mRNA has been deposited in GenBank (accession No. AY561635) and is shown in figure 1. Clone AF024605 has the following variations compared with the other submissions of the same variant. A single nucleotide deletion (G) at position 73 and a nucleotide change from C to T at position 379. Clone NM_145888 (the same variant) has the following nucleotide changes; G to T at position 230; C to A at position 400; G to C at position 1418; A to G at position 505; C to T at position 528; and C to T at position 1238. There is also a single nucleotide insertion (G) at position 1377 (all numbers refer to GenBank accession No. AF024605).

There are two clones that represent *KLK10* splice variant 1 (table 4). Clone BC002710 is 9 bp longer at its 5′ end but is 72 bp shorter than NM_002776 at the 3′ end. The 3′ end of BC002710 is verified by the presence of a poly A tail of 24 nucleotides, whereas the 3′ end of NM_002776 is similar to that of splice variant 2 but does not end with a poly A tail. Likely, these two clones represent two different variants with variable 3′ end. The 5′ end, cannot be conclusively verified for either clone. Clone NM_002776 has the following changes compared to clone BC002710: T replaces G at position 376; A instead of C at position 546; C instead of G at position 564; G instead of A in position 651; T instead of C at position 674, and T instead of C at position 1384 (numbers refer to GenBank Accession No. BC002710). It is worth mentioning that these polymorphisms are similar to those observed in splice variant 2.

Table 5. Summary of the 39 reported SNPs the *KLK10* sequence

3′ UTR SNP		5′ UTR SNP	
ref. SNP ID	nt change	ref. SNP ID	nt change
rs1698	C→A	rs2569454	T→C
rs9524	T→C	rs2691209	A→G
rs10426	T→C	rs3760737	G→A
rs1802056	A→C	rs4417638	T→A
		rs7259651	A→G
Exonic SNPs			
ref. SNP ID	nt change	aa change	aa position
rs3745535	T→G	Ser → Ala	50
rs3097885	G→T	Arg→ Met	145
rs2075690	T→C	Leu → Pro	149
rs1061368	G→C	–	–
rs2075688	C→A	–	–
rs2075689	G→A	–	–
Intronic SNPs			
ref. SNP ID	nt change	refSNP ID	nt change
rs2075687	A→G	rs2569453	C→G
rs2075691	G→A	rs3745536	G→A
rs2075692	A→G	rs6509512	T→C
rs2075693	T→C	rs7250580	A→G
rs2075695	T→C	rs7255834	T→C
rs2075696	C→T	rs7255930	T→C
rs2304157	G→C	rs7259451	G→T
rs2569451	C→T	rs8107630	A→T
rs2569452	G→C		
SNPs flanking the gene			
5′ end		3′ end	
refSNP ID	nt change	refSNP ID	nt change
rs1140280	T→C	rs2739434	A→G
rs1802055	G→A	rs2739433	T→C
rs3760738	G→A	rs7256975	G→A
		rs2569449	C→T

nt = Nucleotide; aa = amino acid.

Two EST clones were identified, AW378611 and AW378618, which have a 3′ extension of the longest and last exon of both variants (exon 6, fig. 1). There is neither a poly A tail nor a polyadenylation signal at the end of these ESTs, thus the exact 3′ extension of these mRNAs cannot be accurately determined.

Single Nucleotide Polymorphisms

As shown in table 5, there are 39 reported SNPs in the *KLK10* sequence. Four SNPs occur in the 3′ UTR of the sequence, and five are located in the 5′ UTR. There are 17 intronic SNPs, and seven SNPs that are found in regions flanking the mRNA sequence of the gene (we denote these as 'locus SNPs'). Three are located in the 5′ end and four in the 3′ end. There are six SNPs throughout the exons of the gene, of which three are predicted to result in an amino acid change. SNP rs2075690 replaces a T with a C which results in an amino acid change from leucine to proline. SNP rs3097885 replaces G with T and changes arginine to methionine. SNP rs3745535 replaces T with a G and results in changing serine to alanine.

Identification of New Splice Variants of KLK10

Comprehensive analysis of 130 EST clones from various databases identified the following new *KLK10* splice variations (fig. 1):

(1) A new splice variant has a different first exon that is located further upstream from the two known variants. We experimentally verified this variant by RT-PCR analysis using variant-specific primers (table 1). No further potential upstream translation start sites were identified in exon 1 for this variant and it is likely that the encoded protein product will be identical to those of SV1 and SV2. The sequence of this variant was deposited in GenBank as *KLK10* splice variant 3 (SV3) (GenBank accession No. AY561634).

(2) Clone AW999510 matches partially with exons 5 and 6. From the middle of exon 6, the last 107 nucleotides of the clone match 100% with a hypothetical protein on chromosome 15q15. It should be noted, however, that this variant might be a chimeric clone that represents a technical artifact. Further experimental verification is, therefore, needed.

(3) Three EST clones AW381905, BE 152152, and AW 381922 have a 5′ extension of exon 3, the longest being a 60-bp extension, ending with a conserved AG splice acceptor site. These clones were found in a head and neck cancer cell line.

(4) Clone AW378611 and AW378618 have a 127 bp 3′ extension of the last exon. There is no poly A tail at the end, so the possibility for further 3′ extension still exists.

(5) An EST clone, AW797746, isolated from a uterus library, has an alternative splice donor site 103 nucleotides upstream in exon 4 with a conserved GT donor site and a new downstream acceptor site of exon 5.

(6) Another clone, BE934269 matches partially in its 5′ end (first 200 nucleotides) with exon 6 of *KLK10*. The last 150 nucleotides match a sequence of a BAC clone of chromosome 4 that partially represents an intronic sequence, continuing as the polycystic kidney disease 2 mRNA. We failed to amplify this variant from a number of randomly selected normal and malignant tissues.

Table 6. In silico analysis of *KLK10* gene expression using the SAGE database

Tissue	Library type	Positivity[a]	Average density (tpm)[b]
Breast	normal	4/8 (50%)	85
	cancer	0/24 (0%)	0
Ovary	normal	1/2 (50%)	20
	cancer	6/11 (66%)	179
Prostate	normal	2/4 (50%)	145
	cancer	3/12 (25%)	30
Stomach	normal	0/1 (0%)	0
	cancer	2/3 (67%)	196
Colon	normal	0/2 (0%)	0
	cancer	1/6 (17%)	131
Pancreas	normal	0/2 (0%)	0
	cancer	4/6 (67%)	52
Skin[c]	unaffected[c]	N/A	605
	actinic keratosis[c]		198
	melanoma		38

N/A = Not applicable.

[a] Defined as the number of libraries with positive gene-specific tags out of the total number of libraries screened.

[b] Tags per million (tpm).

[c] Two libraries were prepared from the same patient with affected and unaffected skin.

(7) BE183486 has a split last exon with an intervening 123 nucleotides of intronic sequence. The splice donor splice is conserved (GT) but the splice acceptor site is not.

(8) Clone BI014258 is identical (100% match) with 300 bp of sequence in Intron II (not shown).

(9) Four clones, BM 353613, BU078698, AI273968 and AW117915 have 95% similarity with another fragment of intron II (not shown).

In silico SAGE Expression Profile

The expression pattern of *KLK10* in normal and cancer libraries, analyzed using the SAGE databases, is summarized in table 6. Our findings clearly verify that *KLK10* is downregulated in breast cancer tissues. *KLK10*-specific tags were detected in 50% of normal breast, with an average expression level of 85 tpm, as compared to no expression in any of 24 cancer libraries. This downregulation is also noted in prostate cancer (average expression of 145 vs. 30 tpm, in normal and cancer libraries, respectively).

The expression level of *KLK10,* on the other hand, is significantly upregulated in ovarian cancer. High expression density (179 tpm) was found in cancer libraries compared to 20 tpm in normal ovarian libraries. Upregulation is also observed in cancers of the gastrointestinal tract. In gastric and colon cancers, high expression levels were found (196 and 131 tpm, respectively) compared to no expression in normal counterparts. The same pattern was observed in pancreatic cancer. Gene-specific tags were detectable in 67% of pancreatic cancer with an average density of 52 tpm.

Comparing two libraries representing the affected and unaffected skin from a patient with actinic keratosis shows a significant drop in expression in the affected, compared to the normal skin (198 vs. 605 tpm, respectively). Low expression levels were seen in melanoma of the skin. Comparable levels of expression were seen in the normal brain and brain tumors (data not shown).

High expression (197 tpm) was also observed in the normal cervix. Moderate expression (85 tpm) was seen in a primary mesothelioma library. Lower levels of *KLK10* are expressed in lung, eye, normal lymph node (20 tpm), and normal white blood cells (31 tpm) (data not shown).

The EST Expression Profile

The EST expression profile of *KLK10* is summarized in table 7. The EST results were consistent with the SAGE data, showing downregulation in breast cancer (25 clones from 5 normal libraries vs. a single clone from one cancer library). Significant upregulation was observed in ovarian cancer (20 clones from 3 cancer libraries, compared to no clones from normal ovary libraries), and endometrial adenocarcinoma (13 clones, all isolated from cancer libraries). The same upregulation pattern was also seen in cancers of the gastrointestinal tract (colon, gastric, pancreatic and esophageal). In the pancreas, two clones were identified from normal libraries, while 18 clones were isolated from cancer tissues. While four clones were identified from the normal prostate, only one clone was found in prostate cancer.

Twenty clones were found in 13 libraries of carcinoma of the head and neck. Other cancers showed few positive clones, including astrocytoma of the brain and squamous cell carcinoma of the skin. No significant difference was observed between normal and cancer libraries in the testis and lung. Few clones were isolated from normal liver, skeletal muscle, and eye.

We experimentally verified the up-regulation of *KLK10* in ovarian cancer by RT-PCR analysis. While the gene was not detectable in the pool of normal ovarian tissues analyzed, a band was noted in seven out of ten cancers (fig. 2). A moderate downregulation was seen in pros-

Table 7. Analysis of *KLK10* expression in the EST databases

Tissue	Library type	Matching clones	Number of positive libraries
Breast	normal	25	5
	carcinoma	1	1
Ovary	normal	0	0
	adenocarcinoma	20	3
Uterus	normal	0	0
	carcinoma	13	5
Colon	normal	0	0
	adenocarcinoma	10	8
Stomach	normal	0	0
	carcinoma	9	3
Pancreas	normal	2	1
	adenocarcinoma	18	2
Esophagus	normal	0	0
	carcinoma	1	1
Prostate	normal	4	2
	carcinoma	1	1
Head and neck	normal	0	0
	carcinoma	20	13
Brain	normal	0	0
	astrocytoma	4	1
Skin	normal	0	0
	carcinoma	2	1
Testis	normal	2	1
	carcinoma	1	1
Lung	normal	5	3
	carcinoma	2	2
Liver	normal	2	1
	carcinoma	0	0
Skeletal muscle	normal	2	1
	carcinoma	0	0
Eye	normal	6	1
	carcinoma	0	0

tate cancer. In four patients, bands of weaker intensity were obtained from the cancerous portion, when compared to the normal counterpart of the same patient. Approximately equal expression was seen in two pairs and higher cancer expression in one pair (data not shown).

Taxonomy Analysis

Taxonomy analysis of the *KLK10* gene is shown in table 8. *KLK10* orthologues were identified in three other species in addition to human. The length of the hK10 polypeptide is comparable in all species with the longest being the human protein. The mouse and rat orthologues showed the highest degree of sequence homology to human (42% each).

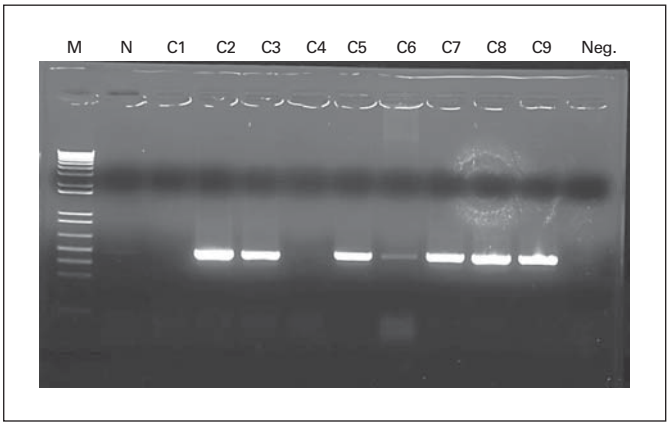

Fig. 2. A representative gel showing *KLK10* expression in normal and cancerous ovarian tissues by RT-PCR. M = Molecular weight marker; N = normal; C = cancer; -ve = negative control. Note high expression in 7/9 cancerous tissues.

Discussion

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 [9]. Gene expression data contained in the electronic databases at the CGAP can be used to identify potentially informative marker genes expressed in cancer. In recent years, these databases are now being used successfully [10–12] and results obtained from in silico analyses have now been experimentally verified in many cases.

The reliability of our analysis was verified by comparing the in silico pattern of kallikrein expression in all normal tissues with previously published reports on kallikrein expression [2, 12, 13]. The results were in general agreement with PCR and Northern blotting data [2, 13–16]. For example, our SAGE and EST data indicate that *KLK2–4* are highly expressed in the prostate, while *KLK5* is highly expressed in normal breast tissue and skin, which is in agreement with published experimental reports [16, 17].

There is significant evidence that *KLK10* is linked to endocrine related malignancies such as ovarian, breast, testicular and prostate cancer with differential expression in cancer compared to normal counterparts [18]. The downregulation of *KLK10* in breast cancer has already been documented by experimental evidence. *KLK10* was

Table 8. Summary of *KLK10* taxonomy results

Species	Name	GenBank accession	Length[a]	Degree of homology	Map location
H. sapiens	KLK10	O43240	276	100%	19q13.4
M. musculus	kallikrein 7; Protease, serine, 6; thymopsin	NP_036002	228	42%	7B2
R. norvegicus	NRPN_RAT neuropsin precursor (NP) (kallikrein 8) (brain serine protease 1)	O88780	225	42%	1q22
D. melanogaster	T13596 trypsin homolog – fruit fly	T13596	227	33%	

[a] Number of amino acids.

first discovered by its dramatic downregulation in a breast cancer cell line [5]. When *KLK10* was transfected into this breast cancer cell line and injected into nude mice, anchorage-independent growth was reduced [7]. *KLK10* mRNA expression decreases during tumor progression [19]. Hypermethylation has been suggested to be responsible for the tumor-specific loss of *KLK10* expression in breast cancer [20].

The in silico ovarian expression data are supported by recently published reports on the possible value of hK10 as a tumor marker for ovarian cancer. Serum hK10 concentrations are significantly elevated in 56% of ovarian cancer patients compared to normal individuals [21]. hK10 concentration was also found to change during ovarian cancer progression. High concentrations of serum hK10 are associated with advanced disease stage, serous histological type, suboptimal debulking, and large residual tumor [22]. Patients with tumors expressing high hK10 levels were more likely to die and relapse when compared to tumors expressing low hK10 levels [22]. Recently, *KLK10* overexpression in ovarian cancer was also detected by microarray analysis [23–25].

An interesting finding is the upregulation of the gene in three different cancers of the gastrointestinal tract. Experimental verification is warranted and the possible diagnostic/prognostic role of *KLK10* in these malignancies should be investigated further. Our data also points to the possible involvement of the *KLK10* gene in the pathogenesis of skin diseases. Two kallikreins, *KLK5* and *KLK7*, are expressed at high concentrations in the epidermis and were shown to be involved in skin diseases [26, 27]. It will be interesting to examine the possible interaction between different kallikreins in the pathogenesis of skin disease.

Splice variants are common among kallikreins. To date, over 50 splice variants have been identified for this family [1, 18]. Characterization of all splice variants for each gene is important. In addition to any potential physiological significance, these variants have the potential of being used for diagnostic applications. Slawin et al. [28] reported the prognostic significance of a splice variant-specific RT-PCR for *KLK2*, in detecting prostate cancer metastasis. Nakamura et al. [29 reported differential expression of the brain and prostate types of *KLK11* between benign, hyperplastic and malignant prostate cancer cell lines.

The presence of *KLK10* orthologues points to the possibility of a preserved function among species. Recent reports indicate that the human and mouse loci are very similar and that the orthologues of the 15 human kallikreins are found in the same order in the syntenic locus (chromosome 7) in the mouse [14, 18, 30, 31]. Similar data were published for the rat [32]. Knockout animal models may provide a useful tool to improve our understanding of the physiological role of kallikreins.

It should be noted that our results are based mainly on database analyses. This in-silico approach has been increasingly and successfully used in recent years. These findings, however, should be experimentally verified. Our analyses provide the foundation for new research directions towards a better understanding of the role of this gene in normal physiology and physiopathology, and its clinical utility.

References

1 Yousef GM, Diamandis EP: Human kallikreins: Common structural features, sequence analysis and evolution. Curr Genomics 2003; 4:147–165.

2 Yousef GM, Diamandis EP: The new human tissue kallikrein gene family: Structure, function, and association to disease. Endocr Rev 2001;22:184–204.

3 Diamandis EP, Yousef GM: Human tissue kallikrein gene family: A rich source of novel disease biomarkers. Expert Rev Mol Diagn 2001; 1:182–190.

4 Diamandis EP, Yousef GM, Clements J, Ashworth LK, Yoshida S, Egelrud T, Nelson PS, Shiosaka S, Little S, Lilja H, Stenman UH, Rittenhouse HG, Wain H: New nomenclature for the human tissue kallikrein gene family. Clin Chem 2000;46:1855–1858.

5 Liu XL, Wazer DE, Watanabe K, Band V: Identification of a novel serine protease-like gene, the expression of which is down-regulated during breast cancer progression. Cancer Res 1996;56:3371–3379.

6 Luo L, Herbrick JA, Scherer SW, Beatty B, Squire J, Diamandis EP: Structural characterization and mapping of the normal epithelial cell-specific 1 gene. Biochem Biophys Res Commun 1998;247:580–586.

7 Goyal J, Smith KM, Cowan JM, Wazer DE, Lee SW, Band V: The role for NES1 serine protease as a novel tumor suppressor. Cancer Res 1998;58:4782–4786.

8 Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.

9 Strausberg RL: The Cancer Genome Anatomy Project: New resources for reading the molecular signatures of cancer. J Pathol 2001;195: 31–40.

10 Nacht M, Ferguson AT, Zhang W, Petroziello JM, Cook BP, Gao YH, Maguire S, Riley D, Coppola G, Landes GM, Madden SL, Sukumar S: Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer. Cancer Res 1999;59:5464–5470.

11 Mitas M, Mikhitarian K, Hoover L, Lockett MA, Kelley L, Hill A, Gillanders WE, Cole DJ: Prostate-Specific Ets (PSE) factor: a novel marker for detection of metastatic breast cancer in axillary lymph nodes. Br J Cancer 2002; 86:899–904.

12 Yousef GM, Polymeris ME, Yacoub GM, Scorilas A, Soosaipillai A, Popalis C, Fracchioli S, Katsaros D, Diamandis EP: Parallel overexpression of seven kallikrein genes in ovarian cancer. Cancer Res 2003;63:2223–2227.

13 Yousef GM, Diamandis EP: Expanded human tissue kallikrein family – a novel panel of cancer biomarkers. Tumour Biol 2002;23:185–192.

14 Yousef GM, Diamandis EP: An overview of the kallikrein gene families in humans and other species: Emerging candidate tumour markers. Clin Biochem 2003;36:443–452.

15 Gan L, Lee I, Smith R, Argonza-Barrett R, Lei H, McCuaig J, Moss P, Paeper B, Wang K: Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region. Gene 2000;257:119–130.

16 Clements J, Hooper J, Dong Y, Harvey T: The expanded human kallikrein (KLK) gene family: genomic organisation, tissue-specific expression and potential functions. Biol Chem 2001;382:5–14.

17 Yousef GM, Diamandis EP: The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation. J Biol Chem 1999;274: 37511–37516.

18 Borgono CA, Michael IP, Diamandis EP: Human tissue kallikreins: Physiologic roles and applications in cancer. Mol Cancer Res 2004; 2:257–280.

19 Dhar S, Bhargava R, Yunes M, Li B, Goyal J, Naber SP, Wazer DE, Band V: Analysis of normal epithelial cell specific-1 (NES1)/kallikrein 10 mRNA expression by in situ hybridization, a novel marker for breast cancer. Clin Cancer Res 2001;7:3393–3398.

20 Li B, Goyal J, Dhar S, Dimri G, Evron E, Sukumar S, Wazer DE, Band V: CpG methylation as a basis for breast tumor-specific loss of NES1/kallikrein 10 expression. Cancer Res 2001;61:8014–8021.

21 Luo L, Bunting P, Scorilas A, Diamandis EP: Human kallikrein 10: a novel tumor marker for ovarian carcinoma? Clin Chim Acta 2001;306: 111–118.

22 Luo LY, Katsaros D, Scorilas A, Fracchioli S, Piccinno R, Rigault de la Longrais IA, Howarth DJ, Diamandis EP: Prognostic value of human kallikrein 10 expression in epithelial ovarian carcinoma. Clin Cancer Res 2001;7: 2372–2379.

23 Adib TR, Henderson S, Perrett C, Hewitt D, Bourmpoulia D, Ledermann J, Boshoff C: Predicting biomarkers for ovarian cancer using gene-expression microarrays. Br J Cancer 2004;90:686–692.

24 Hibbs K, Skubitz KM, Pambuccian SE, Casey RC, Burleson KM, Oegema TR Jr, Thiele JJ, Grindle SM, Bliss RL, Skubitz AP: Differential gene expression in ovarian carcinoma: Identification of potential biomarkers. Am J Pathol 2004;165:397–414.

25 Santin AD, Zhan F, Bellone S, Palmieri M, Cane S, Bignotti E, Anfossi S, Gokden M, Dunn D, Roman JJ, O'Brien TJ, Tian E, Cannon MJ, Shaughnessy J Jr, Pecorelli S: Gene expression profiles in primary ovarian serous papillary tumors and normal ovarian epithelium: Identification of candidate molecular markers for ovarian cancer diagnosis and therapy. Int J Cancer 2004;112:14–25.

26 Caubet C, Jonca N, Brattsand M, Guerrin M, Bernard D, Schmidt R, Egelrud T, Simon M, Serre G: Degradation of corneodesmosome proteins by two serine proteases of the kallikrein family, SCTE/KLK5/hK5 and SCCE/KLK7/hK7. J Invest Dermatol 2004;122: 1235–1244.

27 Komatsu N, Takata M, Otsuki N, Toyama T, Ohka R, Takehara K, Saijoh K: Expression and localization of tissue kallikrein mRNAs in human epidermis and appendages. J Invest Dermatol 2003;121:542–549.

28 Slawin KM, Shariat SF, Nguyen C, Leventis AK, Song W, Kattan MW, Young CY, Tindall DJ, Wheeler TM: Detection of metastatic prostate cancer using a splice variant-specific reverse transcriptase-polymerase chain reaction assay for human glandular kallikrein. Cancer Res 2000;60:7142–7148.

29 Nakamura T, Mitsui S, Okui A, Kominami K, Nomoto T, Miki T, Yamaguchi N: Alternative splicing isoforms of hippostasin (PRSS20/KLK11) in prostate cancer cell lines. Prostate 2001;49:72–78.

30 Olsson AY, Lundwall A: Organization and evolution of the glandular kallikrein locus in *Mus musculus*. Biochem Biophys Res Commun 2002;299:305–311.

31 Diamandis EP, Yousef GM, Olsson AY: An update on human and mouse glandular kallikreins. Clin Biochem 2004;37:258–260.

32 Olsson AY, Lilja H, Lundwall A: Taxon-specific evolution of glandular kallikrein genes and identification of a progenitor of prostate specific antigen. Genomics 2004;84:147–156.