ELSEVIER

Gene 254 (2000) 119–128

GENE
AN INTERNATIONAL JOURNAL ON GENES, GENOMES AND EVOLUTION

www.elsevier.com/locate/gene

The *KLK7* (*PRSS6*) gene, encoding for the stratum corneum chymotryptic enzyme is a new member of the human kallikrein gene family — genomic characterization, mapping, tissue expression and hormonal regulation

George M. Yousef [a,b], Andreas Scorilas [a,b], Angeliki Magklara [a,b], Antoninus Soosaipillai [a], Eleftherios P. Diamandis [a,b,*]

[a] *Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ont. M5G 1X5, Canada*
[b] *Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ont. M5G 1L5, Canada*

Received 14 February 2000; received in revised form 19 May 2000; accepted 19 June 2000
Received by L. Luzzatto

Abstract

The human stratum corneum chymotryptic enzyme (*HSCCE*; *PRSS6*, *KLK7* gene) has been purified from human stratum corneum and is known to participate in the cell shedding process. The cDNA of the gene has previously been reported. Here, we describe the identification of 5′ and 3′ extensions of the published mRNA, and the complete genomic organization of the gene. *KLK7* is composed of five coding exons which have similar lengths to exons of other kallikrein-like genes. The intron phases are completely conserved between this gene and other members of the kallikrein-like gene family. Precise mapping of *KLK7* has indicated that it is located at chromosomal locus 19q13.3–q13.4 between the already known genes *zyme* (*KLK6*) (centromere) and *neuropsin* (*KLK8*) (telomere). Until recently, it was thought that this gene is expressed only in the skin. We here provide evidence that *KLK7* is also expressed at relatively high levels in the central nervous system, kidney, mammary and salivary glands. Its expression is up-regulated by estrogens and glucocorticoids in the breast carcinoma cell line BT-474. The cDNA and protein of this gene are homologous to sequences of other kallikrein-like genes. The gene encodes for a secreted protein. Phylogenetic analysis, the close structural similarities, and its co-localization in the same chromosomal region, suggest that the gene encoding for the stratum corneum chymotryptic enzyme is a new member of the expanded human kallikrein gene family.

Keywords: Cancer biomarkers; Gene structure; Hormonally-regulated genes; Human stratum corneum chymotryptic enzyme; Kallikreins; *KLK7*; *PRSS6*; Serine proteases

1. Introduction

In rodents, the kallikrein gene family consists of more than 20 genes. While some of these are pseudogenes, the rest are highly expressed primarily in the salivary glands (Clements, 1997). In humans, three kallikrein genes have been well characterized, including the pancreatic/renal kallikrein gene (*KLK1*, encoding for hK1 protein), the human glandular kallikrein 2 gene (*KLK2*, encoding for hK2 protein) and prostate-specific antigen (*KLK3*, encoding for hK3 or PSA protein) (Rittenhouse et al., 1998). The latter two human kallikreins have already found important clinical applications as biomarkers for prostatic and breast carcinomas (Rittenhouse et al., 1998; Diamandis, 1998).

More recently, novel serine protease genes which share structural homologies and co-localize in the same chromosomal locus as the classical human kallikreins (chromosome 19q13.3–q13.4) have been cloned (Diamandis et al., 2000). These genes include

Abbreviations: BCM, Baylor College of Medicine; dNTPs, deoxynucleoside triphosphates; EST, expressed sequence tag; HSCCE, human stratum corneum chymotryptic enzyme; HSCTE, human stratum corneum tryptic enzyme; kb, kilobase pairs; KLK, kallikrein; KLK-L, kallikrein-like; LLNL, Lawrence Livermore National Laboratory; *NES1*, normal epithelial cell-specific 1 gene; PCR, polymerase chain reaction; PSA, prostate specific antigen; RT, reverse transcription; TLSP, trypsin-like serine protease.

* Corresponding author. Fax: +1-416-586-8628.
E-mail address: ediamandis@mtsinai.on.ca (E.P. Diamandis)

0378-1119/00/$ - see front matter
PII: S0378-1119(00)00280-8

prostase/*KLK-L1* (*KLK4* according to the approved official kallikrein gene nomenclature) (Nelson et al., 1999), *KLK-L2/HSCTE* (human stratum corneum tryptic enzyme, *KLK5* according to the approved official kallikrein gene nomenclature) (Yousef and Diamandis, 1999; Brattsand and Egelrud, 1999), the *zyme/protease M/neurosin* gene (*PRSS9*, *KLK6*) (Anisowicz et al., 1996; Little, 1997; Yamashiro, 1997), the human stratum corneum chymotryptic enzyme (*HSCCE*; *PRSS6*, *KLK7*) (Hansson et al., 1994), *neuropsin/TADG-14* (*ovasin*; *PRSS19*, *KLK8*) (Shimizu, 1998; Lowell, 1999), the normal epithelial cell-specific 1 gene (*NES1*; *PRSSL1*, *KLK10*) (Goyal et al., 1998), trypsin-like serine protease (*PRSS20*, *KLK11*) (Yoshida et al., 1998), *KLK-L3* (*KLK9*) (Yousef et al., 1999), *KLK-L5* (*KLK12*) (Yousef et al., 1999), *KLK-L4* (*KLK13*) (Yousef et al., 1999) and *KLK-L6* (*KLK14*).

Some of the newly identified kallikrein-like genes have been found to be either underexpressed or overexpressed in certain carcinomas (Diamandis et al., 2000). For example, at the tissue level, *PSA*, *KLK6* and *KLK10* have been reported to be down-regulated while *KLK7* and *KLK8* were found to be highly up-regulated in ovarian carcinomas (Lowell, 1999). Thus, it is possible that a few members of this expanded kallikrein gene family may serve as valuable cancer biomarkers for disease diagnosis or monitoring.

Although the cDNA for the gene encoding human stratum corneum chymotryptic enzyme (*KLK7*) has previously been reported (Hansson et al., 1994), its genomic organization is unknown and its precise mapping, in comparison with other kallikrein genes, has not as yet been elucidated. In this paper, we report the complete genomic organization of *KLK7* and its precise localization, in comparison with other kallikreins. We also study its tissue expression and regulation by steroid hormones, and provide evidence for the presence of alternatively spliced forms of the gene. We further report homology analysis between this gene and other members of the kallikrein gene family. The extensive structural homologies of this gene and protein with other kallikreins, and its chromosomal localization, qualify *KLK7* as a new member of the expanded human kallikrein gene family.

2. Materials and methods

2.1. Construction of an almost contiguous area of DNA sequence on chromosome 19q13.3–q13.4

Large DNA sequencing data for chromosome 19 is available at the web site of the Lawrence Livermore National Laboratory (LLNL) (http://www-bio.llnl.gov/genome/genome. html). We have screened approximately 300 kb of nucleotides around chromosome 19q13.3–q13.4. The sequences were in the form of nine separate contigs of different lengths. We performed a restriction analysis study of the available sequences using the 'WebCutter 2' computer program and with the aid of the EcoR1 restriction map of this area (also available from the LLNL), we were able to construct an almost contiguous area of genomic sequences.

By using the published sequences of KLK3, KLK2, KLK10, KLK1, KLK6, KLK8, KLK11, as well as the software 'BLAST 2 sequences' (Altschul et al., 1997), we were able, using alignment strategies, to identify the relative positions of these genes on the DNA sequence of interest.

2.2. Structure analysis

Multiple alignment was performed using the clustal X software and the multiple alignment program available from the Baylor College of Medicine (BCM) search launcher (kiwi.imgen.bcm.tmc.edu:8808/search-launcher/launcher/html). Phylogenetic studies were performed using the Phylip software package. Distance matrix analysis was performed using the 'Neighbor-Joining/UPGMA' program and parsimony analysis was done using the 'Protpars' program. A hydrophobicity study was performed using the BCM search launcher programs. Signal peptide was predicted using the SignalP software. Protein structure analysis was performed by SAPS program.

2.3. Breast cancer cell line and stimulation experiments

The breast cancer cell line BT-474 was purchased from the American Type Culture Collection (ATCC), Rockville, MD. BT-474 cells were cultured in RPMI media (Gibco BRL, Gaithersburg, MD) supplemented with glutamine (200 mmol/l), bovine insulin (10 mg/l), fetal bovine serum (10%), antibiotics and antimycotics, in plastic flasks, to near confluency. The cells were then aliquoted into 24-well tissue culture plates and cultured to 50% confluency. 24 h before the experiments, the culture media were changed into phenol red-free media containing 10% charcoal-stripped fetal bovine serum. For stimulation experiments, various steroid hormones dissolved in 100% ethanol were added into the culture media, at a final concentration of 10^{-8} M. Cells stimulated with 100% ethanol were included as controls. The cells were cultured for 24 h, then harvested for mRNA extraction.

2.4. Reverse transcriptase polymerase chain reaction

Total RNA was extracted from the breast cancer cells using Trizol reagent (Gibco BRL) following the manufacturer's instructions. RNA concentration was determined spectrophotometrically. 2 μg of total RNA was reverse-transcribed into first strand cDNA using the

Table 1
Primers used for reverse transcription polymerase chain reaction (RT-PCR) analysis

Gene	Primer name	Sequence[a]	Product size (bp)
KLK7	HS1	GAATGAGTACACCGTGCACC	360
	HS2	TGCCAGCGCACAGCATGGAA	
pS2	pS2S	GGTGATCTGCGCCCTGGTCCT	328
	pS2AS	AGGTGTCCGGTGGAGGTGGCA	
PSA	PSAS	TGCGCAAGTTCACCCTCA	754
	PSAAS	CCCTCTCCTTACTTCATCC	
Actin	ACTINS	ACAATGAGCTGCGTGTGGCT	372
	ACTINAS	TCTCCTTAATGTCACGCACGA	

[a] All nucleotide sequences are given in the 5′→3′ orientation.

Superscript™ pre-amplification system (Gibco BRL). The final volume was 20 μl. Two gene-specific primers were designed (see Table 1) and PCR was carried out in a reaction mixture containing 1 μl cDNA, 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μl dNTPs (deoxynucleoside triphosphates), 150 ng of primers and 2.5 units of AmpliTaq Gold DNA polymerase (Roche Molecular Systems, Branchburg, NJ, USA) on a Perkin-Elmer 9600 thermal cycler. The cycling conditions were 94°C for 9 min to activate the Taq Gold DNA polymerase, followed by 43 cycles of 94°C for 30 s, 69°C for 1 min and a final extension at 69°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.

2.5. Tissue expression of KLK7

Total RNA isolated from 28 different human tissues was purchased from Clontech, Palo Alto, CA. We prepared cDNA as described above for the tissue culture experiments and used it for PCR reactions with the primers described in Table 1. Tissue cDNAs were amplified at two dilutions for the *KLK7* reactions (undiluted and 10-fold dilution).

2.6. Cloning and sequencing of the PCR products

To verify the identity of the PCR products, they were cloned into the pCR 2.1-TOPO vector (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The inserts were sequenced from both directions using vector-specific primers, with an automated DNA sequencer.

3. Results

3.1. Chromosomal localization and mapping of the KLK7 gene

Alignment of the known cDNA structure of *KLK7* (Hansson et al., 1994) with the 300 kb genomic sequence allowed us to localize this gene and a number of other known genes, as shown in Fig. 1 (the direction of transcription is denoted by the arrows). The *KLK3* gene is the most centromeric, separated by approximately 119 kb from the *KLK7* gene, which is transcribed in the opposite direction. The *KLK8* gene is 14 kb more telomeric and is transcribed in the same direction as the *KLK7* gene (telomere to centromere). *KLK2*, *KLK4*, *KLK5* and *KLK6* lie between *KLK3* and the *KLK7* gene.

3.2. Genomic structure of the KLK7 gene

The *KLK7* gene, as presented in Fig. 2, is formed of six exons (one non-coding and five coding exons) and five intervening introns, spanning an area of 6 kb of genomic sequence on chromosome 19q13.3–q13.4. The lengths of the coding parts of the exons are 73, 148, 248, 137 and 156 bp, respectively (see our GenBank submission #AF166330). The intron/exon splice sites (GT…AG) and their flanking sequences are closely related to the consensus splicing sites (–mGTAAGT…CAGm–) (Iida, 1990). The protein coding region of the *KLK7* gene is formed of 762 bp nucleotide sequence (including the stop codon), encoding a deduced 253 amino acid polypeptide that will be translated into a protein with a predicted molecular mass of 27.5 kDa. There is one translation initiation codon (ATG) starting at position 59 of the second exon. A possible polyadenylation signal (AATAAA) is present 168 bases downstream of the termination codon.

3.3. Protein product of the KLK7 gene and comparison with the kallikrein protein family

The hydrophobicity plot of *KLK7* shows a hydrophobic region in the N-terminal part of the protein

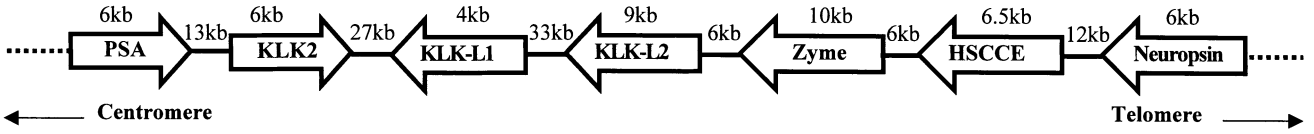

Fig. 1. An approximate 150 kb of contiguous genomic sequence around chromosome 19q13.3–q13.4. Genes are represented by horizontal arrows denoting the direction of transcription. Genomic lengths of genes and distances between genes are shown in kilobases (kb). Figure is not drawn to scale. For full gene names see 'Abbreviations'. The stratum corneum chymotryptic enzyme gene (*HSCCE*) is now also known as *KLK7*.

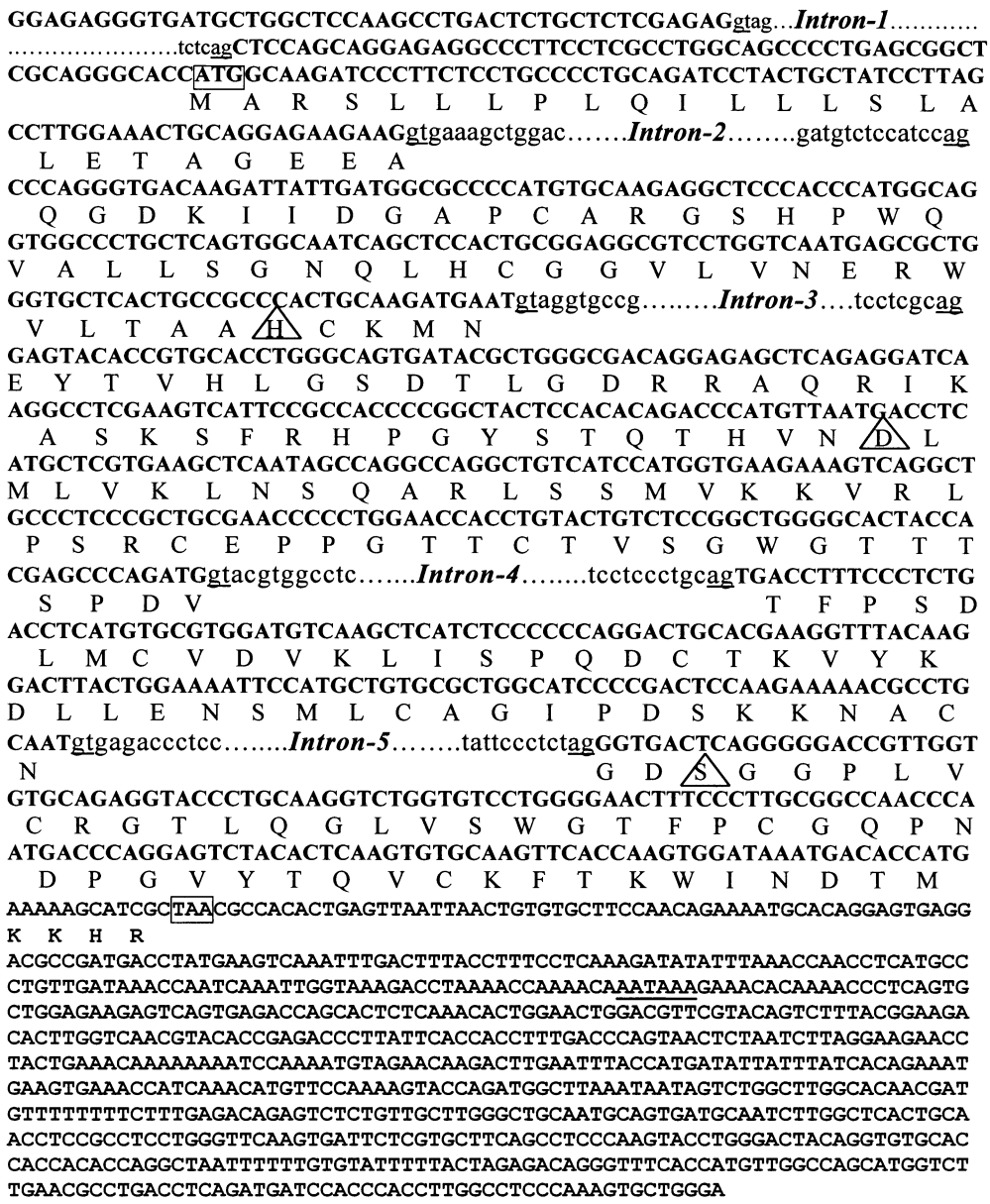

Fig. 2. Genomic organization and partial genomic sequence of the *KLK7* gene. Intronic sequences are not shown except for short areas around the splice junctions. Introns are shown with lowercase letters and exons are in bold and capitalized. The start and stop codons are boxed and the exon–intron junctions are underlined. The catalytic residues are inside triangles. Putative polyadenylation signal is underlined. For full genomic sequence, see our GenBank submission #AF166330. The 3′ end of the mRNA is variable; we describe the longest 3′ end (see also Fig. 5).

(Fig. 3), thus a presumed signal peptide is present. By software analysis, a 22 amino acid signal peptide was predicted with a cleavage site at the carboxyl end of Gly[22]. A potential N-glycosylation site (Asn–Asp–Thr) at position 246 is indicated (Fig. 4). For better characterization of the predicted structural motifs of the *KLK7* protein, it was aligned with other members of the kallikrein multi-gene family (Fig. 4). *KLK7* is found to have 45% amino acid identity and 55% similarity with *KLK8*, 44% identity and 61% similarity with *KLK4*, 43, 42, 40, 39, 39 and 38% identity with *KLK5*, trypsinogen, *KLK11*, hK1, *KLK6* and hK2 proteins, respectively. PSA has 41% identical amino acids with the *KLK7* protein. Multiple alignment study shows that the typical catalytic triad of serine proteases is conserved in the *KLK7* protein (His^{70}, Asp^{112}, and Ser^{205}). As is the case with all other kallikreins, a well-conserved peptide motif is found around the amino acid residues of the catalytic triad (i.e. histidine WVLTAA**H**C, serine GD**S**GGPL, and aspartate **D**IMLV). 12 cysteine residues are present in the mature *KLK7* protein, 10 of them are conserved in all the serine proteases that are aligned in Fig. 4, and would be expected to form disulfide bridges. The other two (C^{137} and C^{239}) are not found in PSA, hK1, hK2

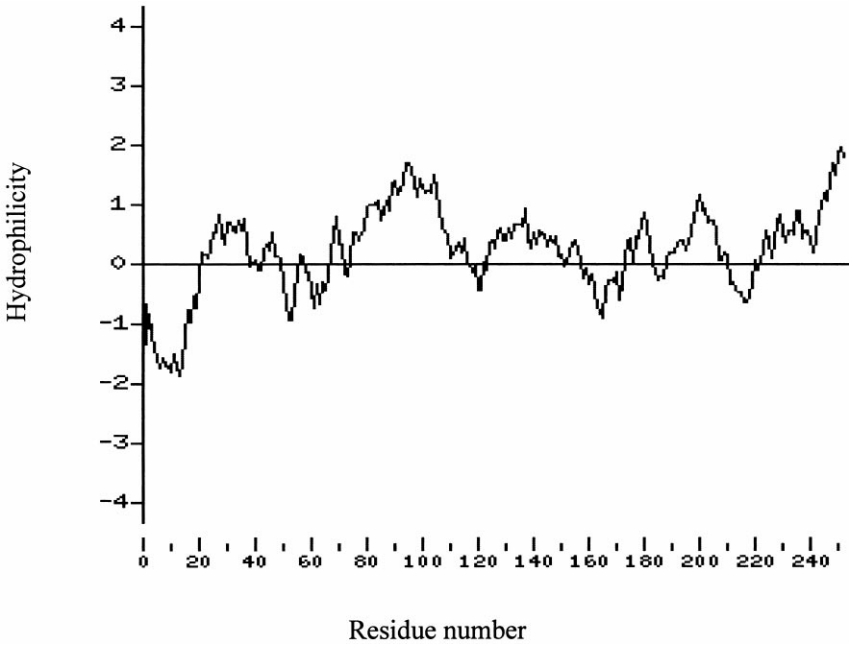

Fig. 3. Plot of hydrophobicity and hydrophilicity of *KLK7* gene product. The hydrophobic region at the beginning of the protein indicates the presence of a signal peptide. For discussion, see text.

and trypsinogen, however, they are found in similar positions in *KLK4*, *KLK11*, *KLK6*, *KLK8*, and *KLK5* proteins and are expected to form an additional disulfide bond. 29 'invariant' amino acids surrounding the active site of serine proteases have been described (Dayhoff, 1978). Of these, 28 are conserved in *KLK7*. One unconserved amino acid (Lys^{169} instead of Pro) is also found in *KLK8* and hK1 proteins. According to protein evolution studies, the changed amino acid represents a conserved evolutionary change to a protein of the same group. The absence of aspartate (D) at position 196 indicates that *KLK7* will not produce a trypsin-like cleavage pattern, like most of the other kallikreins (e.g. hK1, hK2, *KLK8*, *KLK6*, *KLK4*, and *KLK11*). PSA, which has a serine (S) residue in the corresponding position, has chymotrypsin-like activity (Fig. 4). The bold and italic region in Fig. 4 indicates an 10-amino acid loop characteristic of the classical kallikreins (PSA, hK1, and hK2) that is not found in *KLK7* or other members of the kallikrein-like gene family.

3.4. EST homology search

UniGene is an experimental system for automatically partitioning GenBank sequences into a non-redundant set of gene-oriented clusters. In addition to sequences of well-characterized genes, hundreds of thousands of expressed sequence tag (EST) sequences have been included (Schuler et al., 1996). Searching the UniGene database for the *KLK7* gene revealed 16 EST clones with high homology (Table 2). The sequences of these clones were compared with the published mRNA (GenBank accession #L33404) and genomic (GenBank accession #AF166330) sequences of *KLK7*, and all of them show >98% homology with both sequences. Fig. 5 shows a schematic presentation of some of these ESTs. Two clones (GenBank accession #W58737 and AA101043) have 5′ extensions upstream of the published mRNA sequence (Hansson et al., 1994). These extensions show 100% match with the genomic sequence, thus ruling out the possibility of being chimeric clones. Both clones sharply define the same splice acceptor site, on exon 3, as shown in Fig. 5, and thus, the beginning

Table 2
EST clones with >95% homology to exons of *KLK7*

GenBank accession no.	Tissue of origin	IMAGE ID[a]	Sequenced end
AA542994	Ovary	980713	?
W58737	Heart	341508	5′
H39945	Breast	181812	5′
AI139437	Heart	1710172	3′
AA291764	Ovary	724685	5′
AI819857	Lung	2405818	3′
AA101044	Pooled	526892	3′
AI808109	Pooled	2363091	3′
AW088663	Ovary	2393407	3′
AA419284	Ovary	755595	5′
AW177043	Colon	N/A	?
AW081801	Esophagus	2580186	3′
AW206700	Pooled	2722779	3′
AA101043	Pancreas	526892	5′
AW134828	Pooled	2712992	3′
AA425991	Ovary	769539	5′

[a] For IMAGE description, see Lennon et al. (1996).

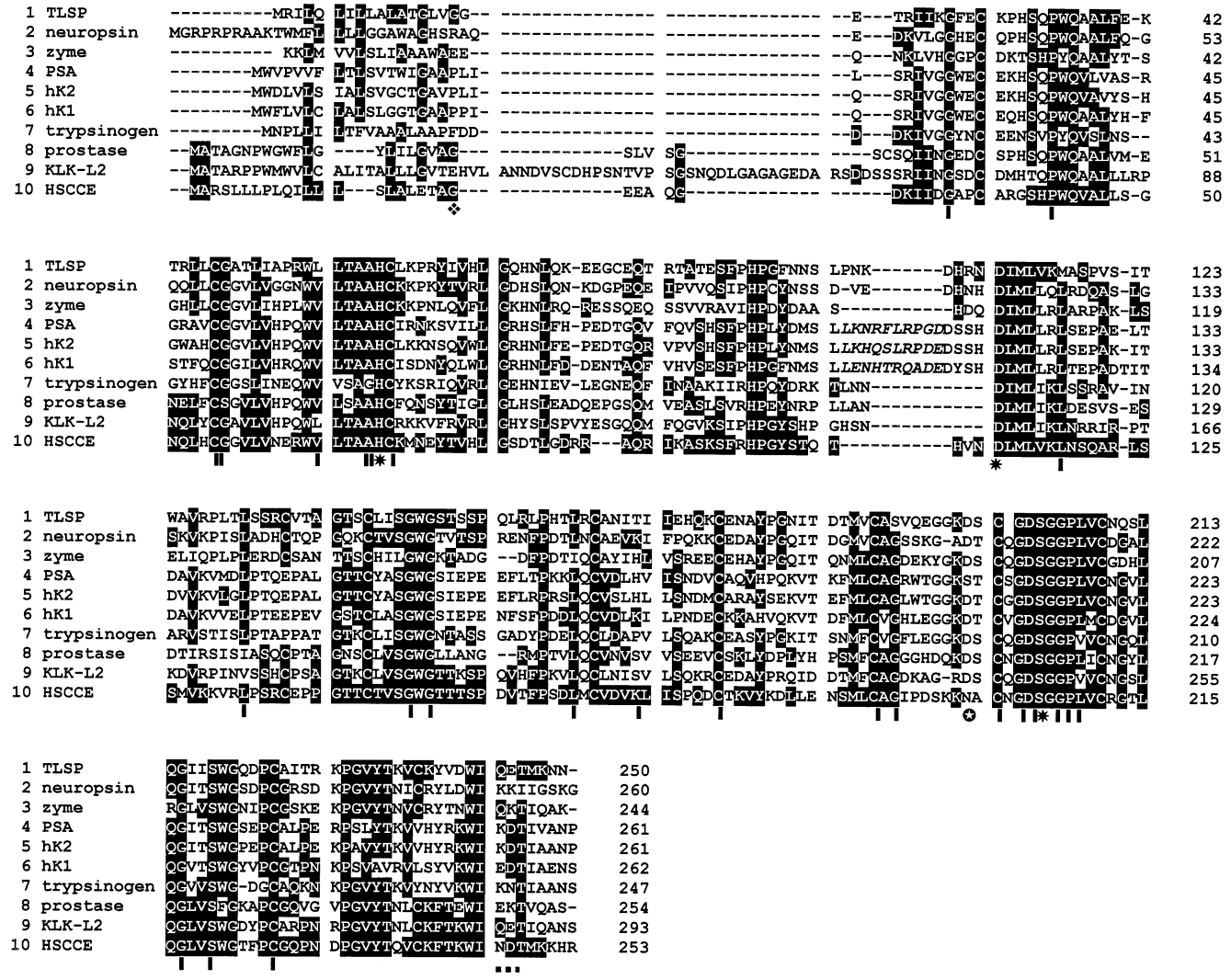

Fig. 4. Alignment of the deduced amino acid sequence of *KLK7* gene product with members of the kallikrein multi-gene family. The residues of the catalytic triad are represented by (✷) and the 29 invariant serine protease residues by (▮ or ✷). The predicted cleavage site is indicated by (❖). The bold and italic area represents the kallikrein loop sequence. The residue D for trypsin-like cleavage pattern is indicated by (✪). *KLK7* has N, and PSA has S in this position.

of the first coding exon should be located 58 bp upstream from the start codon. However, each of the two ESTs has a different 5′ extension upstream of the first coding exon, suggesting the presence of alternatively spliced forms (Fig. 5). The 3′ end of the ESTs extends to a variable distance from the stop codon. Three EST clones (AA101044, AI819857, and AI808109) have a 193 bp extension after the stop codon, in accordance with the published mRNA (Hansson et al., 1994). The longest clone ends 748 bp downstream from the stop codon (this clone is shown in detail in Fig. 2). One of the ESTs (AW134828) has 476 bp of 3′ untranslated region that ends with a polyA tail, suggesting that different polyadenylation signals are utilized by the primary transcript, thus creating mRNAs of different length.

A unique EST clone (accession #AI819857) utilizes a different splice acceptor site for the start of the last exon, which is located upstream from the splice junction utilized by all other ESTs. Thus, this gene may have a number of different splice variants, as shown in Fig. 5.

3.5. Tissue expression of the KLK7 gene

Tissue expression of the *KLK7* gene was assessed by RT-PCR. Actin was used as a control and *PSA* was used as another control with tissue restricted specificity. The *PSA* gene was found to be highly expressed in the prostate, as expected, and to a lower extent in mammary, thyroid and salivary glands, as described in recent literature reports (Diamandis, 1998). The tissue expression of *KLK7* is summarized in Fig. 6. This protease is primarily expressed in the brain, mammary gland, cerebellum, spinal cord, kidney and skin. Lower levels are seen in salivary glands, uterus, thymus, thyroid, pla-

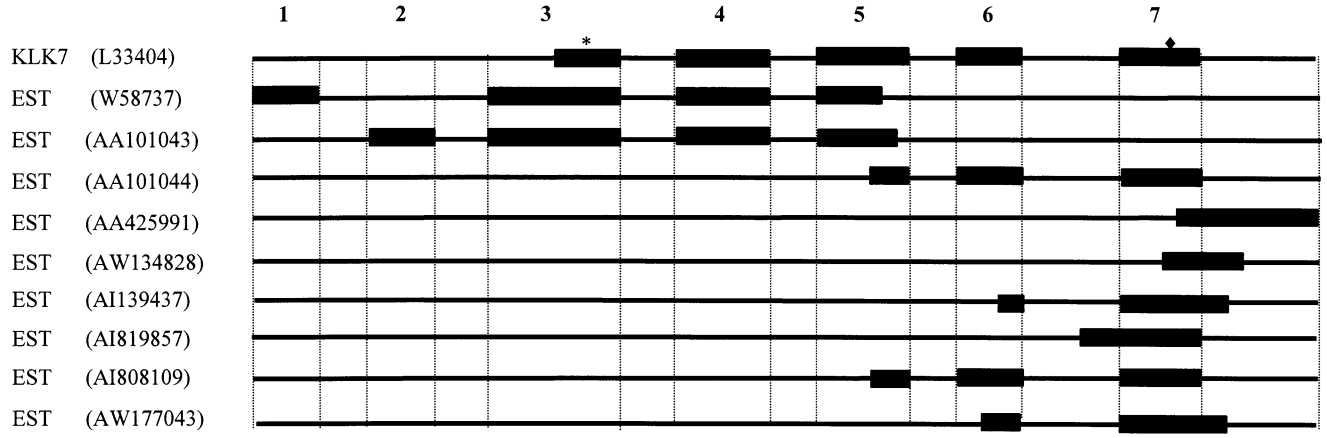

Fig. 5. A schematic presentation comparing the mRNA sequence of the *KLK7* gene with some EST clones. Exons are presented by solid bars and introns by the connecting lines. Top row represents the published mRNA sequence (18), and underneath are the EST clones with the GenBank accession numbers in brackets. Exons are numbered according to our GenBank submission (AF166330). The position of the start codon is denoted by *, and of the stop codon by ♦.

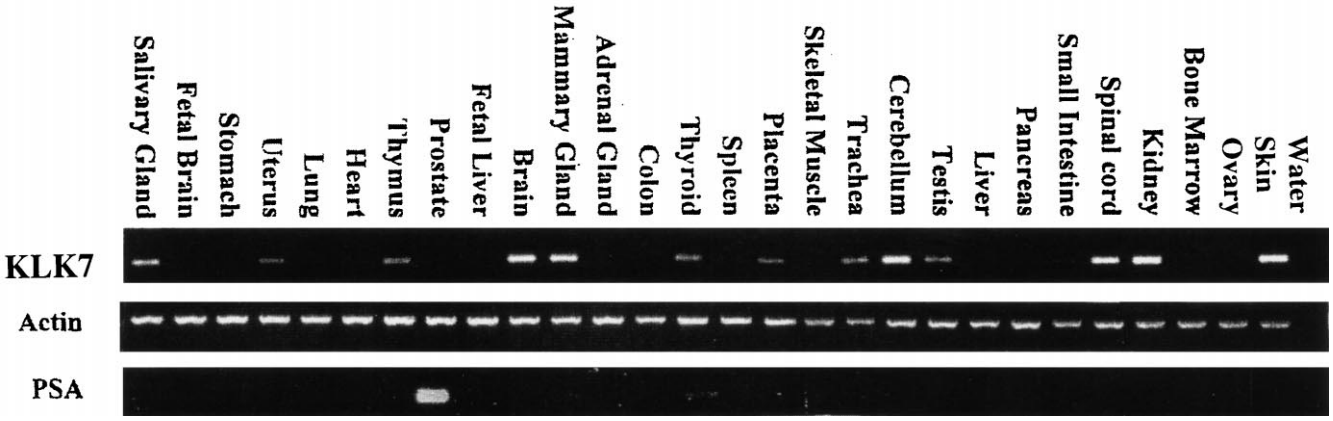

Fig. 6. Tissue expression of the *KLK7* gene as determined by RT-PCR. Actin and PSA are control genes. For discussion, see text.

centa, trachea and testis. The other tissues tested were negative. In order to verify the RT-PCR specificity, the PCR products were cloned and sequenced.

3.6. Hormonal regulation of KLK7 gene

We have used the steroid hormone receptor-positive breast carcinoma cell line BT-474 as a model system to preliminarily examine whether the *KLK7* gene is under steroid hormonal regulation. *PSA* was used as a control known to be up-regulated by androgens and progestins and *pS2* as an estrogen up-regulated control. The *KLK7* gene was found to be up-regulated primarily by estrogens and glucocorticoids. This up-regulation was dose-responsive and it was evident at steroid hormone levels $\geq 10^{-10}$ M (data not shown).

3.7. Phytogenetic analysis

Fig. 7 shows that the *KLK7* is grouped with another two kallikrein-like proteins (*KLK5*, *KLK4*), whereas the classical kallikreins and other serine proteases were separated in other groups.

4. Discussion

The stratum corneum chymotryptic enzyme was first identified as a chymotrypsin-like proteinase that may be involved in the desquamation process in plantar stratum corneum (Lundstrom and Egelrud, 1991). The enzyme has been shown to be secreted to the stratum corneum extracellular space and catalyze the degradation of desmosomes during remodeling of the deeper layers of the skin. Although the cDNA of this serine protease has already been cloned (Hansson et al., 1994), its genomic organization has not been studied. We undertook the task of studying the genomic structure of this gene and compared it with structures of already known serine proteases and other human kallikrein-like genes.

The genomic organization, precise mapping at the chromosomal locus 19q13.3–q13.4, and other attributes

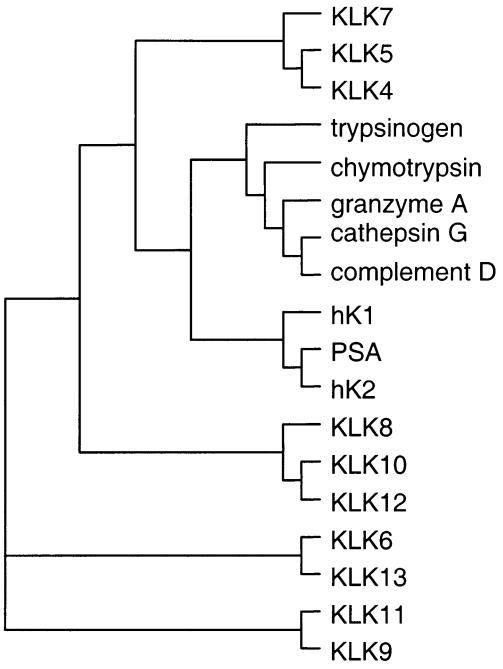

Fig. 7. Dendrogram of the predicted phylogenetic tree for some kallikrein proteins. Neighbor-joining/UPGMA method was used to align *KLK7* with other members of the kallikrein gene family. The tree grouped the classical kallikreins (hK1, hK2, and PSA) together and aligned the *KLK7* protein in one group with *KLK4* and *KLK5*.

of this gene allow us to classify it as a new member of the expanded human kallikrein gene family (Fig. 1) (Diamandis et al., 2000). Our proposal is based on the following observations. (a) The gene maps in the same chromosomal locus as other serine proteases of the kallikrein gene family. More precisely, the stratum corneum chymotryptic enzyme gene lies between the *KLK6* gene (centromere) and *KLK8* (telomere). (b) The gene is composed of five coding exons which have similar lengths to other members of the kallikrein gene family (Fig. 8). (c) The intron phases are completely conserved between this gene and all other members of the kallikrein gene family, as partially shown in Fig. 8. (d) The cDNA and amino acid sequence of the gene show significant homologies with other kallikreins (Fig. 4). (e) Phylogenitically, the stratum corneum chymotryptic enzyme is closely related to *KLK5* and *KLK4*, two newly identified kallikrein-like proteins (Fig. 7). This relationship between *KLK7*, *KLK5* and *KLK4* was confirmed with different methods for phylogenetic analysis. This is despite the localization of these genes on the chromosomal locus. *KLK7* lies adjacent to *KLK6* (centromere) and *KLK8* (telomere), while *KLK5* and *KLK4* both map centromerically to *KLK6*. (f) The stratum corneum chymotryptic enzyme has a hydrophobic signal peptide and a putative activation site which are similar to other members of the kallikrein gene family. These data

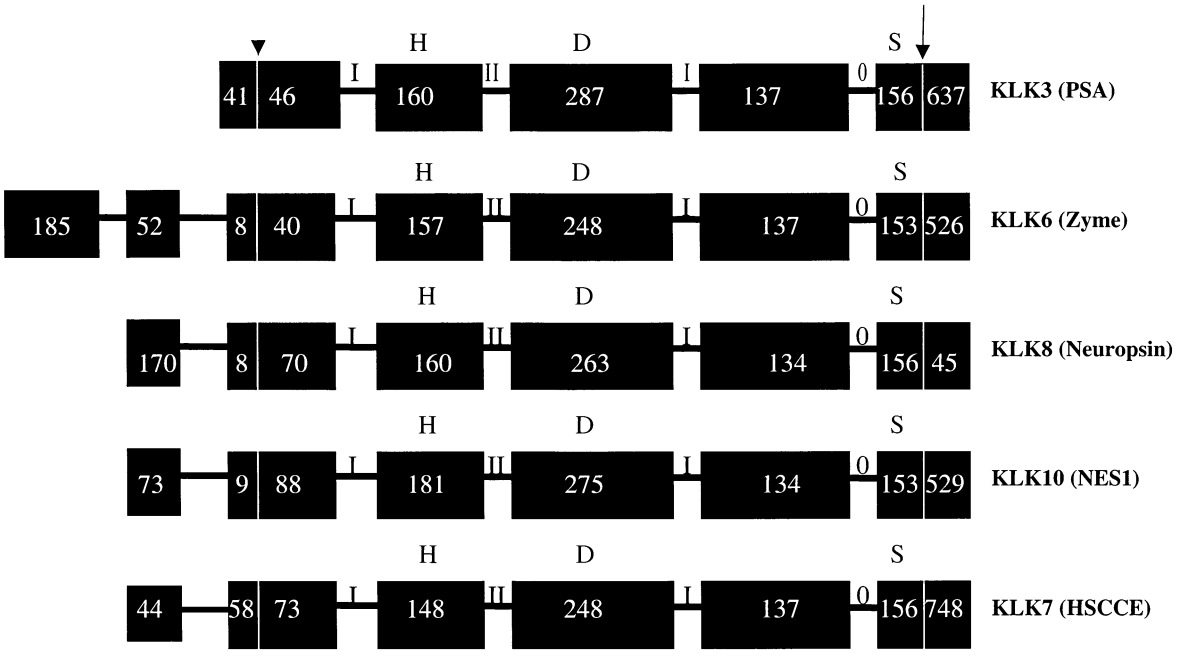

Fig. 8. Schematic diagram showing the comparison of the genomic structure of *PSA* (*KLK3*), *KLK6* (*zyme*), *KLK8* (*neuropsin*), *KLK10* (*NES1*) and *KLK7* (*HSCCE*) genes. Exons are shown by black boxes and introns by the connecting lines. The arrowhead shows the start codon, and the arrow the stop codon. Letters above boxes indicate the relative positions of the amino acids of the catalytic triad; H denotes histidine, D aspartic acid, and S serine. Roman numbers indicate intron phases. The intron phase refers to the location of the intron within the codon; I denotes that the intron occurs after the first nucleotide of the codon, II the intron occurs after the second nucleotide, 0 the intron occurs between codons. Numbers inside boxes indicate exon lengths in base pairs. Figure is not drawn to scale.

suggest that the *KLK7* protein is secreted, consistent with already published information (Hansson et al., 1994; Lundstrom and Egelrud, 1991). We have further preliminarily shown, in parallel to findings for many other kallikrein-like genes (Diamandis et al., 2000), that the *KLK7* gene is regulated by steroid hormones in model breast carcinoma cell line systems. This regulation was evident at steroid hormone concentrations of 10^{-8}–10^{-10} M, which is roughly the physiological range of these hormones in serum. These data suggest that the regulation of *KLK7* by steroid hormones is likely to be relevant in vivo.

Previous reports indicate that this gene is primarily expressed in the skin and to a lower extent in other tissues (Lundstrom and Egelrud, 1991; Ekholm and Egelrud, 1999). By using RT-PCR, we found that the gene is expressed at relatively high amounts in a number of tissues, including kidney, the central nervous system (spinal cord, cerebellum and brain) as well as in mammary and salivary glands. These results were verified by sequencing the PCR products, and by finding EST clones isolated from a variety of tissues (Table 2). The wide range of tissue expression of *KLK7* should not be surprising. By using RT-PCR instead of Northern blot analysis, many kallikrein genes were found to be expressed in a wide variety of tissues, including salivary gland, kidney, pancreas, brain, and tissues of the reproductive system (uterus, mammary gland, ovary, and testis) (Diamandis et al., 2000).

Our results suggest the existence of alternatively spliced forms and other variants of the *KLK7* gene (Fig. 5). A previous study, using Northern blot analysis, showed the presence of two mRNA transcripts which differ in the 3′ untranslated region (Hansson et al., 1994). A number of kallikrein genes express multiple alternatively spliced forms (Heuze et al., 1994; Riegman et al., 1988; Riegman et al., 1991; Chen et al., 1994; Mitsui et al., 1999).

The data presented here strongly suggest that the *KLK7* gene is a new member of the expanded human kallikrein gene family. This family likely originated from an ancestral gene-by-gene duplication. Recent data further indicate that, in addition to its involvement in certain skin diseases (Ekholm and Egelrud, 1999), this enzyme may also be involved in carcinogenesis. Tanimoto et al. (1999) have demonstrated overexpression of this serine protease in a subset of ovarian carcinomas. Recently, many other kallikrein-like genes have been linked to carcinogenesis, including *PSA*, *KLK2*, *KLK10*, *KLK8*, *KLK6*, and *KLK9* (reviewed by Diamandis et al., 2000).

We conclude that the stratum corneum chymotryptic enzyme is a new kallikrein-like protein which may find applications as a biomarker in skin diseases as well as in cancer. It will be interesting to study the protein levels of this enzyme in the general circulation for the purpose of diagnosing or monitoring malignant and non-malignant disease.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST, a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Anisowicz, A., Sotiropoulou, G., Stenman, G., Mok, S.C., Sager, R., 1996. A novel protease homolog differentially expressed in breast and ovarian cancer. Mol. Med. 2, 624–636.

Brattsand, M., Egelrud, T., 1999. Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation. J. Biol. Chem. 274, 30 033–30 030.

Chen, L.M., Murray, S.R., Chai, K.X., Chao, L., Chao, J., 1994. Molecular cloning and characterization of a novel kallikrein transcript in colon and its distribution in human tissues. Brazilian J. Med. Biol. Res. 27, 1829–1838.

Clements, J.A., 1997. The molecular biology of the kallikreins and their roles in inflammation. In: Farmer, S. (Ed.), The Kinin System. Academic Press, New York, pp. 71–79.

Dayhoff, M.O., 1978. Atlas of protein sequence and structure. Natl. Biomed. Res. Found. 5, 79–81.

Diamandis, E.P., 1998. Prostate-specific antigen — its usefulness in clinical medicine. Trends Endocrinol. Metab. 9, 310–316.

Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A., Obiezu, C., 2000. The new kallikrein gene family — implications in carcinogenesis. Trends Endocrinol. Metab. 11, 54–60.

Ekholm, E., Egelrud, T., 1999. Stratum corneum chymotryptic enzyme in psoriasis. Arch. Dermatol. Res. 29, 195–200.

Goyal, J., Smith, K.M., Cowan, J.M., Wazer, D.E., Lee, S.W., Band, V., 1998. The role for NES1 serine protease as a novel tumor suppressor. Cancer Res. 58, 4782–4786.

Hansson, L., Stromqvist, M., Backman, A., Wallbrandt, P., Carlstein, A., Egelrud, T., 1994. Cloning, expression and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase. J. Biol. Chem. 269, 19 420–19 426.

Heuze, N., Olayat, S., Gutman, N., Zani, M., Courty, Y., 1994. Molecular cloning and human tissues. Brazilian J. Med. Biol. Res. 27, 1829–1838.

Iida, Y., 1990. Quantification analysis of 5′-splice signal sequences in mRNA precursors. Mutations in 5′-splice signal sequence of human beta-globin gene and beta-thalassemia. J. Theor. Biol. 145, 523–533.

Lennon, G., Auffray, C., Polymeropoulos, M., Soares, M.B., 1996. The IMAGE consortium: an integrated molecular analysis of genomes and their expression. Genomics 33, 151–152.

Little, S.P., 1997. Zyme, a novel and potentially amyloidogenic enzyme cDNA isolated from Alzheimer's disease brain. J. Biol. Chem. 272, 5135–5142.

Lowell, J., 1999. Cloning of tumor-associated differentially expressed gene-14, a novel serine protease overexpressed by ovarian carcinoma. Cancer Res. 59, 4435–4439.

Lundstrom, A., Egelrud, T., 1991. Stratum corneum chymotryptic enzyme: a proteinase which may be generally present in the stratum corneum and with a possible involvement in desquamation. Acta Derm. Venereol. 71, 471–474.

Mitsui, S., Tsuruoka, N., Yamashiro, K., Nakazato, H., Yamaguchi, N., 1999. A novel form of human neuropsin, a brain-related serine protease, is generated by alternativesplicing and is expressed preferentially in human adult brain. Eur. J. Biochem. 260, 627–634.

Nelson, P.S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L.,

Wang, K., 1999. Molecular cloning and characterization of prostase, an androgen-regulated serine protease with prostate-restricted expression. Proc. Natl. Acad. Sci. USA 96, 3114–3119.

Riegman, P.H., Klaassen, P., van der Korput, J.A., Romijn, J.C., Trapman, J., 1988. Molecular cloning and characterization of novel prostate antigen cDNAs. Biochem. Biophys. Res. Commun. 155, 181–188.

Riegman, P.H., Vlietstra, R.J., van der Korput, H.A., Romijn, J., Trapman, J., 1991. Identification and androgen-regulated expression of two major human glandular kallikrein-1 (hGK-1) mRNA species. Mol. Cell Endocrinol. 76, 181–190.

Rittenhouse, H.G., Finlay, J.A., Mikolajczyk, S.D., Partin, A.W., 1998. Human kallikrein 2 (hK2) and prostate-specific antigen (PSA), two closely related, but distinct, kallikreins in the prostate. Crit. Rev. Clin. Lab. Sci. 35, 275–368.

Schuler, G.D., Boguski, M.S., Stewart, E.A., Stein, L.D., Gyapay, G., Rice, K., White, R.E., 1996. A gene map of the human genome. Science 274, 540–546.

Shimizu, C., 1998. Characterization of recombinant and brain neuropsin, a plasticity-related serine protease. J. Biol. Chem. 273, 11 189–11 196.

Tanimoto, H., Underwood, L.J., Shigemasa, K., Yan, M.S.Y., Clarke, J., Parmley, T.H., O'Brien, T.J., 1999. The stratum corneum chymotryptic enzyme that mediates shedding and desquamation of skin cells is highly overexpressed in ovarian tumor cells. Cancer 86, 2074–2082.

Yamashiro, K., 1997. Molecular cloning of a novel trypsin-like serine protease (neurosin) referentially expressed in brain. Biochim. Biophys. Acta 1350, 11–14.

Yoshida, S., Taniguchi, M., Suemoto, T., Oka, T., He, X., Shiosaka, S., 1998. cDNA cloning and expression of a novel serine protease, TSLP. Biochim. Biophys. Acta 1399, 225–228.

Yousef, G., Diamandis, E.P., 1999. The new kallikrein-like gene KLK-L2: molecular characterization, mapping, tissue expression and hormonal regulation. J. Biol. Chem. 274, 37 511–37 516.

Yousef, G.M., Luo, L.-Y., Diamandis, E.P., 1999. Identification of novel human kallikrein-like genes on chromosome 19q13.3–q13.4. Anticancer Res. 19, 2843–2852.