Genomics **69,** 331–341 (2000)
doi:10.1006/geno.2000.6346, available online at http://www.idealibrary.com on IDEAL®

KLK12 Is a Novel Serine Protease and a New Member of the Human Kallikrein Gene Family—Differential Expression in Breast Cancer

George M. Yousef, Angeliki Magklara, and Eleftherios P. Diamandis[1]

Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, M5G 1X5, Canada; and Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, M5G 1L5 Canada

Received May 18, 2000; accepted August 7, 2000

Kallikreins are a subgroup of serine proteases that are involved in the posttranslational processing of polypeptide precursors. Growing evidence suggests that many kallikreins are implicated in carcinogenesis. In rodents, kallikreins are encoded by a large multigene family, but in humans, only three genes have been identified. By using the positional candidate approach, we were able to identify a new kallikrein-like gene, tentatively named KLK12 (for kallikrein gene 12). This new gene maps to chromosome 19q13.3–q13.4, is formed of five coding exons, and shows structural similarity to serine proteases and other known kallikreins. KLK12 is expressed in a variety of tissues including salivary gland, stomach, uterus, lung, thymus, prostate, colon, brain, breast, thyroid, and trachea. We identified three splicing forms of KLK12 that are expressed in many tissues. Our preliminary results indicate that the expression of KLK12 is down-regulated at the mRNA level in breast cancer tissues and is up-regulated by steroid hormones in breast and prostate cancer cell lines. This gene may be involved in the pathogenesis and/or progression of certain cancer types and may find applicability as a novel cancer biomarker.

INTRODUCTION

The progressive development of genetic tools, together with the availability of powerful computer programs and the large amount of sequence information generated by the Human Genome Project, has enabled the mapping of several loci responsible for human diseases, including cancer, as well as the identification of areas where certain gene families are clustered. Positional candidate cloning is a relatively new approach for gene discovery that combines the knowledge of map position with the increasingly dense human transcript maps, the availability of expressed sequence tags (ESTs),[2] and the identification of candidate genes mapped to the same region (Ballabio, 1993). This method greatly expedites the search process and is already one of the predominant methods of gene discovery (Collins, 1995).

Kallikreins (KLKs) are a subfamily of serine proteases, originally defined as enzymes cleaving vasoactive peptides (kinins) from kininogen (Bhoola *et al.*, 1992; Schachter, 1979). This enzyme family consists primarily of plasma kallikrein and tissue or glandular kallikreins. Plasma kallikrein is encoded by a single gene that is structurally different from genes encoding tissue kallikrein (Clements, 1997). The tissue or glandular kallikreins are involved in posttranslational modification of polypeptides and are crucial to many biological processes (Bhoola *et al.*, 1992; Clements, 1997; Rittenhouse *et al.*, 1998). The human tissue kallikrein gene family was until recently thought to consist of only three members: the pancreatic-renal kallikrein (KLK1) (Evans *et al.*, 1988), the glandular kallikrein (KLK2) (Schedlich *et al.*, 1987), and prostate-specific antigen (PSA or KLK3) (Riegman *et al.*, 1989). The introduction of prostate-specific antigen (PSA) testing about 15 years ago has revolutionized the management of patients with prostate cancer (Diamandis, 1998). The diagnostic usefulness of PSA in prostate cancer led us to speculate that other related molecules might be valuable biomarkers of different malignancies, including breast and ovarian cancer. In addition to PSA, hK2 (human glandular kallikrein 2; encoded by the KLK2 gene) has been proposed as an adjuvant diagnostic marker for prostate cancer (Partin

[1] To whom correspondence and reprint requests should be addressed at Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. Telephone: (416) 586-8443. Fax: (416) 586-8628. E-mail: ediamandis@mtsinai.on.ca.

[2] Abbreviations used: KLK, kallikrein; KLK-L, kallikrein-like; PCR, polymerase chain reaction; PSA, prostate-specific antigen; hK2, human glandular kallikrein 2; BCM, Baylor College of Medicine; EST, expressed sequence tag; EMSP1, enamel matrix serine proteinase 1; TLSP, trypsin-like serine protease; HSCCE, human stratum corneum chymotryptic enzyme; NES1, normal epithelial cell-specific 1 gene; RACE, rapid amplification of cDNA ends; ER, endoplasmic reticulum.

0888-7543/00 $35.00

et al., 1999; Stenman, 1999). Moreover, accumulating evidence indicates that other members of the expanded kallikrein gene family may be associated with malignancy (Diamandis *et al.*, 2000a).

In our efforts to identify new kallikrein-like genes that might be involved in malignancy, we have examined a 300-kb genomic region on chromosome 19q13.3–q13.4 that contains the kallikrein gene locus. We were able to discover a new gene, tentatively named KLK-L5 (for kallikrein-like gene 5). The Human Gene Nomenclature Committee-approved symbol for this gene is KLK12, and we will use this symbol throughout this article (Diamandis *et al.*, 2000b). Here, we describe the cloning of the new gene, its genomic and mRNA structure, its precise location in relation to other known kallikreins, and its tissue expression pattern. We also describe the cloning of three different splice variants of the KLK12 gene. Our preliminary data suggest that this gene is down-regulated in breast cancer and that it is under steroid hormone regulation in breast and prostate cancer cell lines.

MATERIALS AND METHODS

DNA sequence on chromosome 19. We have obtained sequencing data of approximately 300 kb of nucleotides on chromosome 19q13.3–q13.4 from the Web site of the Lawrence Livermore National Laboratory (http://www-bio.llnl.gov/genome/genome.html). This sequence was in the form of nine contigs of different lengths. Restriction enzyme analysis and long polymerase chain reaction (PCR) strategies, followed by DNA sequencing, bacterial artificial chromosome and P1-derived artificial chromosome library screening, and end sequencing of selected clones, were used to construct a contiguous genomic region, representing the complete human kallikrein gene locus (our unpublished data; see also below).

New gene identification. A number of computer programs were used to predict the presence of putative new genes within the contiguous genomic area of interest. The ability of these programs to predict new genes was first examined by using the genomic sequences of the known kallikreins as testing parameters. The most reliable computer programs, GeneBuilder (gene prediction) (http://l25.itba.mi.cnr.it/~webgene/genebuilder.html), GeneBuilder (exon prediction) (http://l25.itba.mi.cnr.it/~webgene/genebuilder.html), Grail 2 (http://compbio.ornl.gov), and GENEID-3 (http://apolo.imim.es/geneid.html), were selected for further use.

Expressed sequence tag searching. The predicted exons of the putative new gene were subjected to homology search using the BLASTN algorithm (Altschul *et al.*, 1997) on the National Center for Biotechnology Information Web server (http://www ncbi.nlm.nih.gov/BLAST/) against the human EST database. A clone with >95% homology was obtained from the IMAGE consortium (Lennon *et al.*, 1996) through Research Genetics Inc. (Huntsville, AL). This clone was propagated, purified, and sequenced from both directions with an automated sequencer, using insert-flanking vector primers.

Rapid amplification of cDNA ends (RACE). According to the EST sequence and the predicted structure of the gene, two sets of gene-specific primers were designed for 5′ and 3′ RACE reactions. Two rounds of RACE reactions (nested PCR) were performed for each type of RACE with 5 μl Marathon Ready cDNA of human testis and prostate (Clontech, Palo Alto, CA) as templates. The reaction mix and PCR conditions were selected according to the manufacturer's recommendations. In brief, the initial denaturation was for 5 min at 94°C, followed by 94°C for 5 s and 72°C for 2 min, for 5 cycles; then 94°C for 5 s and 70°C for 2 min, for 5 cycles; and then 94°C for 5 s and 65°C for 2 min for 30 cycles for the first reaction and 25 cycles for the nested PCR.

Tissue expression. Total RNA isolated from 26 different human tissues was purchased from Clontech. We prepared cDNA as described below for the tissue culture experiments and used it for PCRs. After all known kallikrein genes were aligned, two primers (KLK12-R1 and KLK12-F1) (Table 1 and Fig. 1) were designed from areas with relatively low homology.

Tissue cDNAs were amplified at various dilutions. Due to the high degree of homology between kallikreins, and to exclude nonspecific amplification, PCR products were cloned and sequenced.

Normal and malignant breast tissues. Normal breast tissues were obtained from women undergoing reduction mammoplasties. Breast tumor tissues were obtained from female patients at participating hospitals of the Ontario Provincial Steroid Hormone Receptor Program. The normal and tumor tissues were immediately frozen in liquid nitrogen after surgical resection and stored in this manner until extracted. The tissues were pulverized with a hammer under liquid nitrogen, and RNA was extracted as described below, using Trizol reagent.

Breast and prostate cancer cell lines and hormonal stimulation experiments. The breast cancer cell lines BT-474 and T-47D and the LNCaP prostate cancer cell line were purchased from the American Type Culture Collection (Rockville, MD). Cells were cultured in RPMI medium (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. Twenty-four hours before the experiments, the culture medium was changed into phenol red-free medium containing 10% charcoal-stripped fetal bovine serum. For stimulation experiments, various steroid hormones dissolved in 100% ethanol were added into the culture medium at a final concentration of 10^{-8} M. Cells stimulated with 100% ethanol were included as controls. The cells were cultured for 24 h and then harvested for mRNA extraction.

Reverse transcriptase-polymerase chain reaction (RT-PCR). Total RNA was extracted from the cell lines or tissues using Trizol reagent (Gibco BRL) following the manufacturer's instructions. RNA concentration was determined spectrophotometrically. Two micrograms of total RNA was reverse-transcribed into first-strand cDNA using the Superscript preamplification system (Gibco BRL). The final volume was 20 μl. Based on the combined information obtained from the predicted genomic structure of the new gene and the EST sequences, two gene-specific primers were designed (KLK12-F1 and KLK12-R1) (Table 1), and PCR was carried out in a reaction mixture containing 1 μl of cDNA, 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM deoxynucleoside triphosphates, 150 ng of primers, and 2.5 units of AmpliTaq Gold DNA polymerase (Roche Molecular Systems, Branchburg, NJ) 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 and 63°C for 1 min, and a final extension step at 63°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, Carlsbad, CA) according to the manufacturer's instructions. The inserts were sequenced from both directions using vector-specific primers, with an automated DNA sequencer.

Structure analysis. Multiple alignment was performed using the Clustal X software package available at ftp://ftp.ebi.ac.uk/pub/software/dos/clustalw/clustalx/(clustalx1.64b.msw.exe) and the multiple alignment program available from the Baylor College of Medicine (BCM; Houston, TX) (kiwi.imgen.bcm.tmc. edu:8808/search-launcher/launcher/html). Phylogenetic studies were performed using the Phylip software package available at http://evolution.genetics.

TABLE 1

Primers Used for RT-PCR Analysis

Gene	Primer name	Sequence[a]
KLK12	KLK12-F1	TCAGCCAGGCAGCCACACCG
	KLK12-R1	TTGGTGATGCCCCAGCCTGA
	KLK12-F2	CCACACCGAAGATTTTCAAT
	KLK12-R2	GCCCCTCCTTCATTTATA
PSA	PSAS	TGCGCAAGTTCACCCTCA
	PSAAS	CCCTCTCCTTACTTCATCC
Actin	ACTINS	ACAATGAGCTGCGTGTGGCT
	ACTINAS	TCTCCTTAATGTCACGCACGA

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

washington.edu/phylip/getme.html. Distance matrix analysis was performed using the "Neighbor-Joining/UPGMA" program, and parsimony analysis was performed using the "Protpars" program. A hydrophobicity study was performed using the BCM search launcher program (http://dot.imgen.bcm.tmc.edu:9331/seq-search/struc-predict.html). The signal peptide was predicted using the "SignalP" server (http://www.cbs.dtu.dk/services/signal). Protein structure analysis was performed by the "SAPS" (structural analysis of protein sequence) program (http://dot.imgen.bcm.tmc.edu:9331/seq-search/struc-predict.html).

RESULTS

Identification of the KLK12 Gene

Computer analysis of the genomic area of interest (300 kb around chromosome 19q13.3–q13.4) predicted a putative gene composed of at least three exons. Screening of the human EST database revealed an EST clone (GenBank Accession No. AI394679) with 99% homology to our predicted exons. This clone was obtained, purified, and sequenced. The full-length sequence of the EST was compared with the genomic area containing the putative new gene and showed 100% homology with certain areas (exons), which were separated by introns. This alignment indicated that the new gene was composed of seven exons. Sequence homology comparisons and phylogenetic analysis revealed that this new gene is structurally similar to known kallikreins and other serine proteases (see below). Since four other new kallikrein-like genes were discovered in this area (Diamandis *et al.*, 2000a; Yousef *et al.*, 1999a, 1999b, 2000; Yousef and Diamandis 1999, 2000), this gene was named KLK-L5 (for kallikrein-like gene 5). Now this gene is known as KLK12 (Diamandis *et al.*, 2000b). Attempts to translate the coding region in all three possible reading frames indicated that only one reading frame will produce a full-length polypeptide chain without interrupting in-frame stop codons. Further support for the correctness of this reading frame was obtained by demonstrating that only this frame will preserve the three amino acid residues necessary for serine protease activity (catalytic triad) and the conserved motifs around them. An in-frame methionine start codon was found in the second exon. This start codon falls within a typical consensus Kozak sequence (CCACC**ATG**G) (Kozak, 1991). Thus, the gene will have at least one 5′ untranslated exon, similar to other kallikrein-like genes [e.g., zyme (Little *et al.*, 1997), the normal epithelial cell-specific 1 gene (NES1) (Luo *et al.*, 1998), and neuropsin (Yoshida *et al.*, 1998)]. 5′ and 3′ RACE reactions were performed to obtain the 5′ and 3′ ends of the gene. No more sequence was obtained by 5′ RACE. However, 3′ RACE enabled us to identify the 3′ untranslated region of the gene. The additional sequence ends with a poly(A) stretch that does not exist in the genomic structure, thus marking the 3′ end of the gene and the start of the poly(A) tail.

Splice Variants of the KLK12 Gene

Homology analysis of the KLK12 gene with other kallikreins revealed the presence of an additional 3′ exon, an observation that has not been reported for any other member of the kallikrein multigene family. Furthermore, two different PCR bands were obtained with the 3′ RACE. Sequencing of these bands revealed that this gene has at least two splice forms at its 3′ end: one form in which the last exon is a single continuous fragment, and another form in which the last exon is split into two exons, with an intervening intron. To identify the full structure of other possible splice variants of the gene, PCR was performed using two primers (L5-F2 and L5-R2) (Table 1 and Fig. 1). We used cDNA from 26 different tissues as templates, and the reaction was performed under different experimental conditions (annealing temperature, $MgCl_2$ concentra-

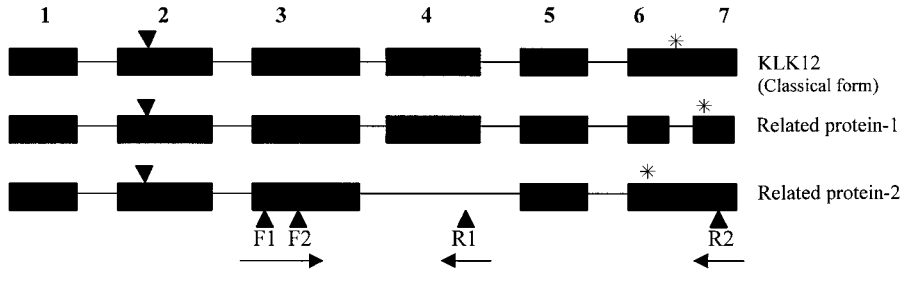

FIG. 1. Diagram showing the comparative genomic structure of the three splice forms of KLK12; the classic kallikrein form, related protein-1, and related protein-2. Exons are represented by solid bars and introns by the connecting lines. Exon numbers refer to Genbank Accession No. AF135025. Start codons are represented by the inverted arrowhead (▼), and stop codons are represented by asterisks (*). Primer locations are represented by vertical arrowheads (▲) and their directions by horizontal arrows. For primer sequences and codes, see Table 1. For more details, see text. Figure is not drawn to scale.

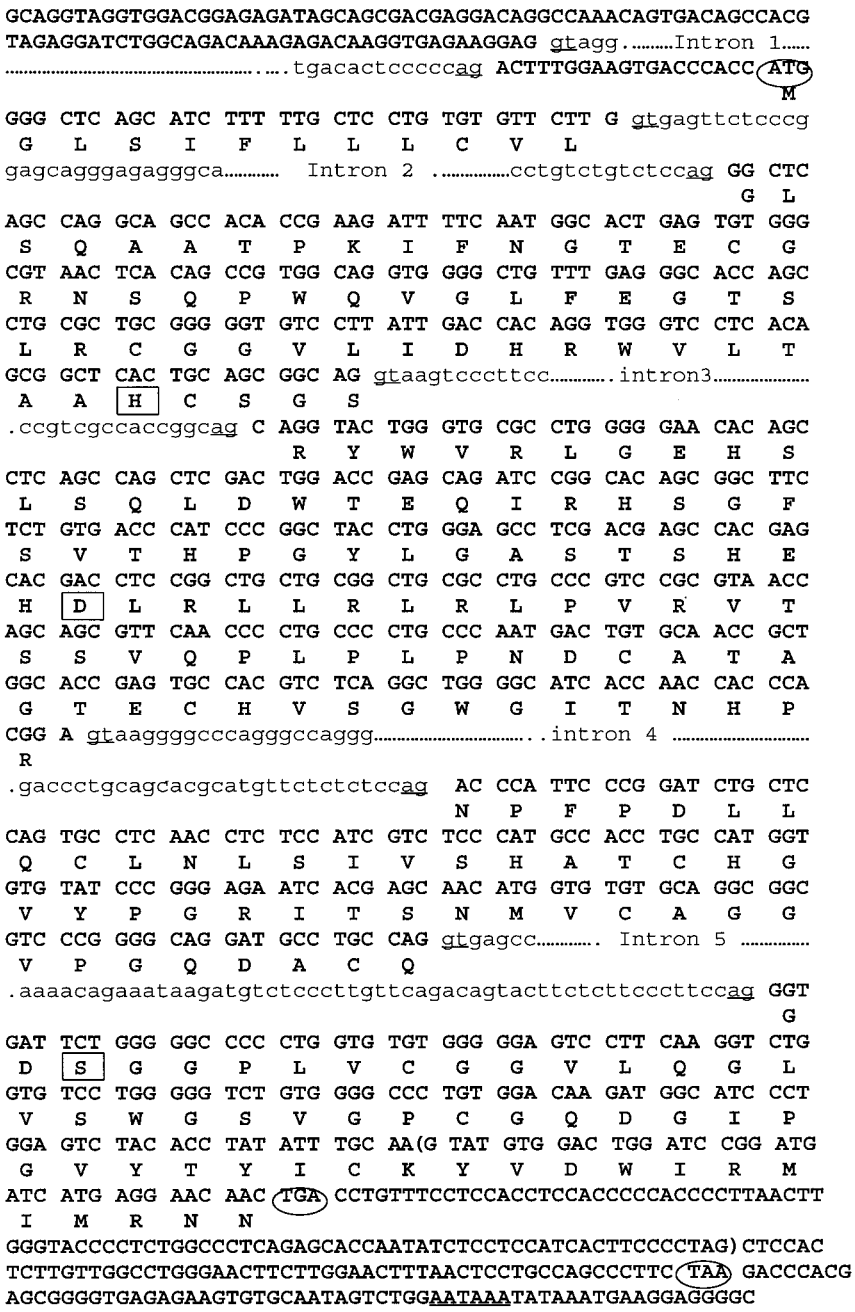

FIG. 2. Genomic organization and partial genomic sequence of the KLK12 gene. Intronic sequences are not shown except for short sequences around the splice junctions. Introns are shown with lowercase letters and exons with uppercase letters. For full sequence, see GenBank Accession No. AF135025. The start and stop codons are encircled, and the exon–intron junctions are underlined. The translated amino acids of the coding region are shown underneath by a single-letter abbreviation. The catalytic residues are boxed. Putative polyadenylation signal is underlined. The extra intron of the related protein-1 form is represented by nonboldface capital letters in parentheses. When this intron is spliced, the frame continues with codon AAC (asparagine, N, instead of lysine, K) until it encounters the stop codon TAA (encircled). For more details, see text.

tion). Three distinct bands were observed in many tissues. These bands were excised, gel-purified, and sequenced. As shown in Fig. 1, the KLK12 gene was found to have three molecular forms: (1) One form (referred to, from now on, as the "classical" form) represents a typical kallikrein-like serine protease with five coding exons and four intervening introns (Fig. 1). As is the case with some other kallikreins, a 5′ untranslated exon is also present, and the possibility of a further upstream untranslated exon(s) could not be excluded. Exons 1, 2, and 3 were present at the aforementioned EST. The start codon is present in the second exon (numbers refer to our GenBank submission, Accession No. AF135025). The stop codon is located in the sixth exon, followed by a 3′ untranslated region, and a typical polyadenylation signal (AATAAA) is located 16 bp before the poly(A) tail (Fig. 2). This form of KLK12 spans a genomic length of 5801 bp on chromo-

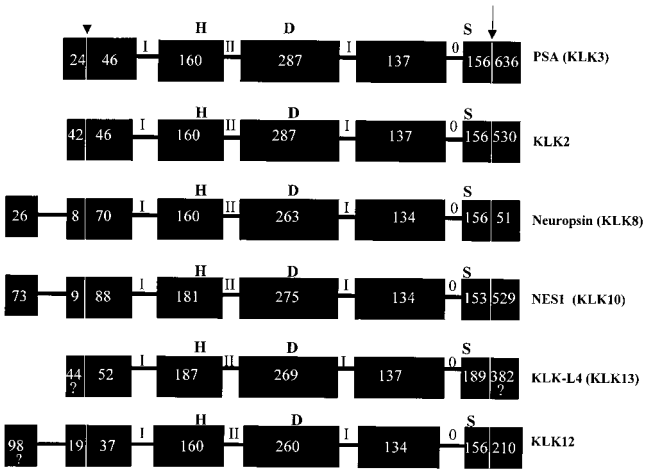

FIG. 3. Schematic diagram showing the comparison of the genomic structure of PSA (KLK3), KLK2, neuropsin (KLK8), NES1 (KLK10), KLK-L4 (KLK13), and KLK12 genes. Exons are shown by solid bars and introns by the connecting lines. Arrowhead marks the site of the start codon, and the arrow represents the stop codon. Letters above boxes indicate relative positions of the catalytic triad; H denotes histidine, D denotes aspartic acid, and S denotes serine. Roman numbers indicate intron phases. The intron phase refers to the location of the intron within the codon: I indicates that the intron occurs after the first nucleotide of the codon, II indicates that the intron occurs after the second nucleotide, and 0 indicates that the intron occurs between codons. Numbers inside boxes indicate exon lengths in basepairs. Question marks indicate that exact length is not accurately known. Figure is not drawn to scale.

some 19q13.3–q13.4. The lengths of the coding regions of the exons are 37, 160, 260, 134, and 156 bp, respectively (Figs. 2 and 3). The predicted protein-coding region is formed of 747 bp, encoding a deduced 248-amino-acid protein with a predicted molecular mass of 26.7 kDa. The intron/exon splice sites (GT . . . AG) and their flanking sequences are in agreement with the consensus splice site sequence (Iida, 1990). (2) The second mRNA form, encoding the KLK12-related protein-1, is an alternatively spliced form in which the last exon is split into two separate exons with an additional intervening intron (Fig. 1). This splitting of the last exon results in the utilization of another stop codon at position 9478, thus creating a deduced 254-amino-acid protein that is 6 amino acids longer than the "classical" KLK12 form and its carboxy-terminal end is different in sequence by 19 amino acids (Fig. 2). This variant has a predicted molecular mass of 27.1 kDa (for base numbering please see our GenBank submission, Accession No. AF135025). (3) The third mRNA form, encoding KLK12-related protein-2, is similar to the classical form except that the fourth exon is missing (Fig. 1). This leads to frameshifting of the coding region, and an earlier in-frame stop codon will be encountered at position 9180. The protein-coding region of this form consists of 336 bp, encoding a predicted 111-amino-acid protein with a molecular mass of 12 kDa. This protein will lack both the serine and the aspartate residues characteristic of serine proteases.

Structural Analysis of the Classical KLK12 Gene

Figure 4 shows a comparative hydrophobicity analysis of the KLK12 and the PSA proteins. The amino-terminal regions of both genes are quite hydrophobic, indicating that this region of KLK12 possibly harbors a signal peptide analogous to PSA. We predicted a cleavage site for a signal peptide between amino acids 17 and 18 (A ↓ A) and an activation peptide between amino acids 21 and 22 (K ↓ I). Figure 4 also shows the presence of several evenly distributed hydrophobic regions throughout the KLK12 polypeptide, which are consistent with a globular protein, similar to other serine proteases (Liu *et al.*, 1996). In Fig. 5, we present alignment of the KLK12 protein with another 10 members of the same gene family. The dotted region in Fig. 5 indicates an 11-amino-acid loop characteristic of the classical human kallikreins (PSA, hK1, and hK2) but not found in KLK12 protein or other members of the kallikrein multigene family (Little *et al.*, 1997; Liu *et al.*, 1996; Yoshida *et al.*, 1998; Yousef and Diamandis 1999, 2000; Yousef *et al.*, 1999b). Sequence analysis of eukaryotic serine proteases indicates the presence of 29 invariant amino acids (Dayhoff, 1978). Twenty-eight of them are conserved in the KLK12 polypeptide, and the remaining amino acid (S^{156} instead of P) is not conserved among all other kallikreins (Fig. 5). Twelve cysteine residues are present in the putative mature KLK12 protein, 10 of them are conserved in all kallikreins, and the remaining 2 (C^{133} and C^{235}) are present in most of the other kallikrein-like proteins but not in the classical kallikreins, and they are expected to form an additional disulfide bridge (Fig. 5).

The presence of aspartate (D) at position 194 suggests that KLK12 will possess a trypsin-like cleavage pattern, similar to most of the other kallikreins [e.g.,

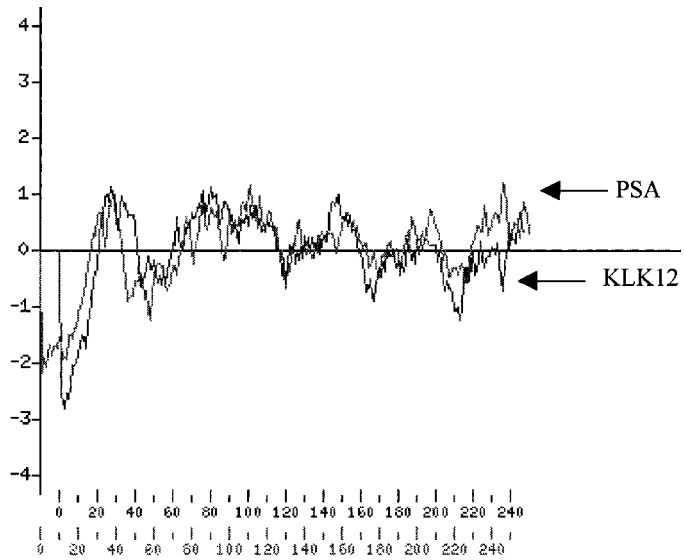

FIG. 4. Plot of hydrophobicity and hydrophilicity of KLK12 protein compared to PSA. The hydrophobic N-terminus may harbor a signal and activation peptide. For details, see text.

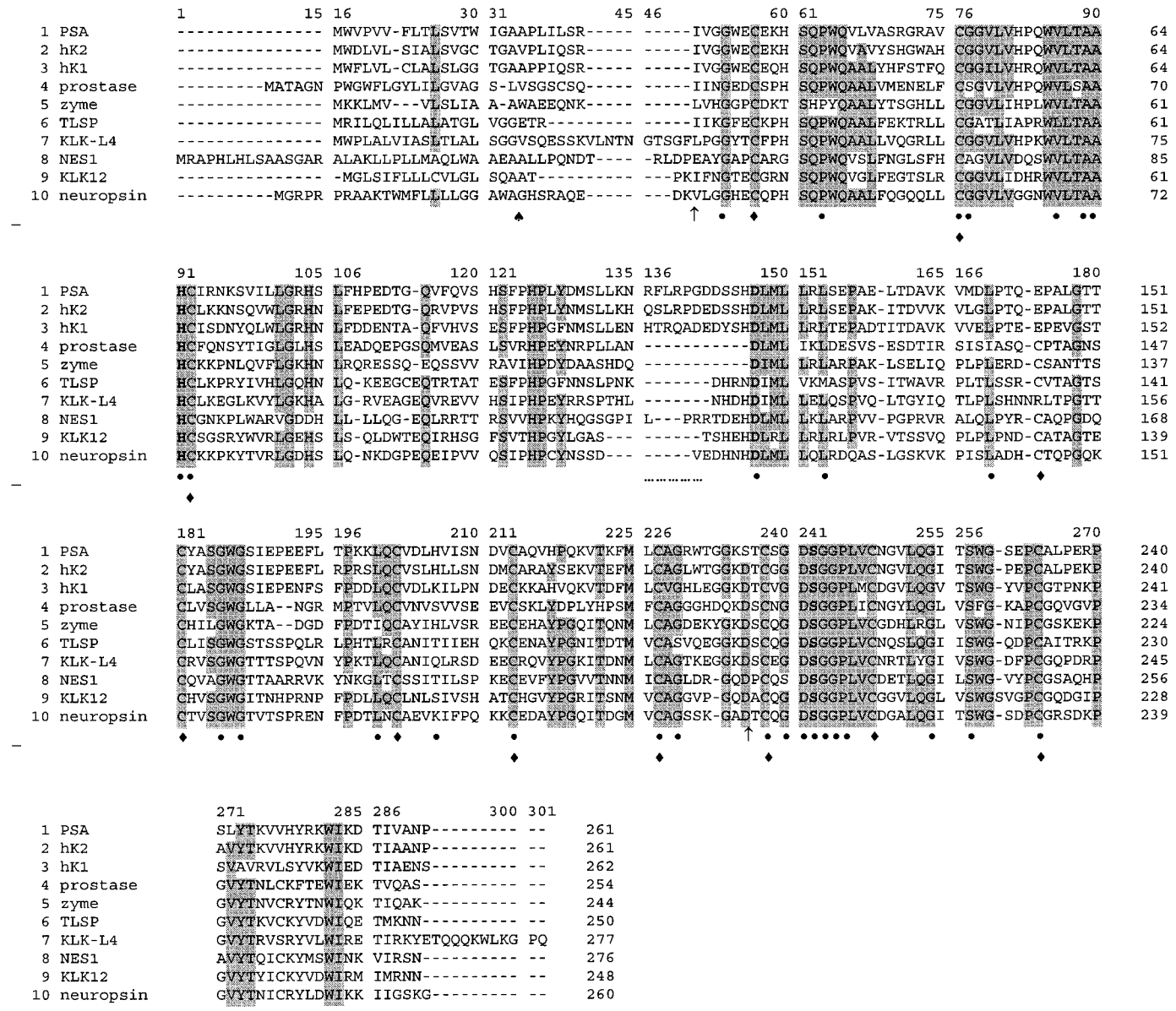

FIG. 5. Alignment of the deduced amino acid sequence of KLK12 with members of the kallikrein multigene family. For full gene names, see abbreviations footnote. According to the new nomenclature (Diamandis *et al.*, 2000b), gene symbols are as follows: KLK3 (PSA), KLK2 (hK2), KLK1 (hK1), KLK4 (prostase), KLK6 (zyme), KLK11 (TLSP), KLK13 (KLK-L4), KLK10 (NES1), and KLK1 (neuropsin). Dashes represent gaps to bring the sequences to better alignment. The residues of the catalytic triad are represented by boldface letters, and the 29 invariant serine protease residues are marked with (•). The cysteine residues are marked by (◆). Conserved areas are highlighted in gray. The predicted cleavage sites in signal peptide are indicated by (♠). The predicted site for the activation peptide is indicated by (↑). The dotted area represents the kallikrein loop sequence. A vertical arrow marks the trypsin-like cleavage site.

hK1, hK2, trypsin-like serine protease (TLSP), neuropsin, zyme, prostase, and EMSP] but different from PSA, which has a serine (S) residue in the corresponding position and is known to have chymotrypsin-like activity (Fig. 5) (Rittenhouse *et al.*, 1998).

Homology with Other Members of the Kallikrein Multigene Family

Although the protein encoded by the KLK12 gene is unique, it has a high degree of homology with the other kallikrein-like genes. The KLK12 protein (classical form) has 48% amino acid sequence identity and 57% overall similarity with neuropsin, 46% identity with NES1, and 38% identity with both PSA and hK2 proteins. Multiple alignment shows that the typical catalytic triad of serine proteases is conserved in the KLK12 protein (H^{62}, D^{108}, and S^{200}) (Figs. 2 and 5). In addition, a well-conserved peptide motif is found around the amino acid residues of the catalytic triad as is the case with other serine proteases [i.e., histidine (VLTAA*H*C), serine (GD*S*GGP), and aspartate (*D*L-RLL)] (Little *et al.*, 1997; Yamashiro *et al.*, 1997) (Fig.

5). Figure 5 also shows other amino acid residues that are completely conserved between kallikreins and kallikrein-like proteins. To predict the phylogenetic relatedness of the KLK12 protein with other serine proteases, the amino acid sequences of the kallikrein proteins were aligned together using the "Clustal X" multiple alignment program, and a distance matrix tree was predicted using the neighbor-joining/UPGMA and Protpars parsimony methods. Figure 6 shows the phylogenetic analysis, which separated the classical kallikreins (hK1, hK2, and PSA) and clustered KLK12 with NES1 and neuropsin proteins in a separate group away from other serine proteases, consistent with previously published studies (Irwin *et al.*, 1988; Nelson *et al.*, 1999) and indicating that this group of genes probably arose from a common ancestral gene, by gene duplication.

Tissue Expression and Hormonal Regulation of the KLK12 Gene

As shown in Fig. 7, the KLK12 gene is expressed primarily in the salivary gland, stomach, uterus trachea, prostate, thymus, lung, colon, brain, breast, and thyroid gland, but, as is the case with other kallikreins, lower levels of expression are found in some other tissues (testis, pancreas, small intestine, and spinal cord). To verify the RT-PCR specificity, the PCR products were cloned and sequenced. The three splice forms of the gene were expressed in most of these tissues. However, the relative abundance of each form was different among tissues (Fig. 7).

To investigate whether the KLK12 gene is under steroid hormone regulation, two breast cancer cell lines (BT-474 and T-47D) and a prostate cancer cell line (LNCaP) were used as models. In LNCaP cells, the gene was up-regulated only by androgen and progestin. In only this cell line were all three isoforms were expressed. In BT-474 cells, KLK12 was found to be up-regulated, at the mRNA level, by estrogen and androgen and to a lesser extent by the progestin. The rank of potency was estrogen > androgen > progestin. However, the rank of potency for the T-47D cell line was androgen > progestin > estrogen. In both of these cell lines, only the short isoform (related protein-2) was present (Fig. 8). To verify the accuracy and specificity of the stimulation experiments, two control genes were used: PSA (upregulated by androgens and progestins) and pS2 (up-regulated by estrogens). The results obtained for the genes were as expected.

KLK12 Is Down-regulated in Breast Cancer

Expression of the KLK12 gene, at the mRNA level, was compared among 17 breast cancer tissues and 1 normal breast tissue, by RT-PCR. The gene is not expressed at all in 12 tumors (Fig. 9). In all breast tissues (normal and malignant), the short isoform (related protein-2) was predominant, with the exception of one tumor, which expressed only the classical form (Fig. 9, lane 8). These results should be interpreted as preliminary, since the number of tumors and normal tissues tested is relatively small.

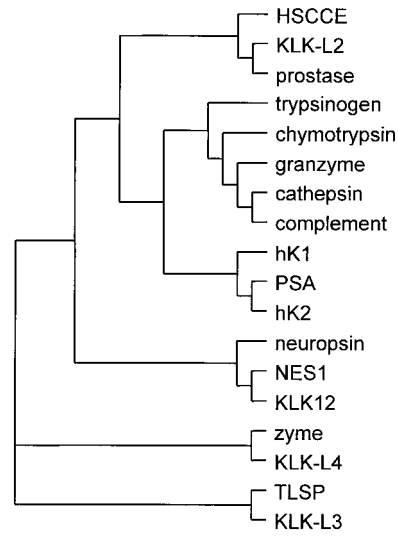

FIG. 6. Dendrogram of the predicted phylogenetic tree for some serine proteases and other kallikrein proteins. Neighbor-joining/UPGMA method was used to align KLK12 with other serine proteases and members of the kallikrein gene family. The tree grouped the classical kallikreins (hK1, hK2, and PSA) together and aligned the KLK12 protein in one group with NES1 and neuropsin. Other serine proteases were aligned in different groups.

Mapping and Chromosomal Localization of the KLK12 Gene

The knowledge of extensive genomic sequence on chromosome 19q13.3–q13.4 enabled us to localize 14 kallikrein-like genes precisely and to determine the direction of transcription, as shown by the arrows in Fig. 10. Only PSA and KLK2 transcribe from centromere to telomere; the rest of the genes are transcribed in the reverse direction. The KLK1 gene was found to be the most centromeric, and the KLK-L6 gene (also known as KLK14) was found to be the most telomeric (KLK-L6; GenBank Accession No. AF161221). KLK12 is 21.3 kb centromeric to KLK13 (GenBank Accession No. AF135024) and 1.6 kb more telomeric to the TLSP gene (Fig. 10).

DISCUSSION

Kallikreins play important roles in diverse physiological processes, and this action is linked to their serine protease activity (Bhoola *et al.*, 1992; Clements, 1997). A number of structural criteria have been proposed to classify putative kallikreins into this multigene family (Irwin *et al.*, 1988). As shown in Fig. 3, kallikreins are characterized by the following common structural features (Diamandis *et al.*, 2000a): (a) All genes are formed of five coding exons and four intervening introns [some genes may have an extra 5′ untranslated exon(s)] (Luo *et al.*, 1998; Yoshida *et al.*, 1998; Yousef *et al.*, 1999a). (b) The exon lengths are usually comparable. (c) The intron phases are always

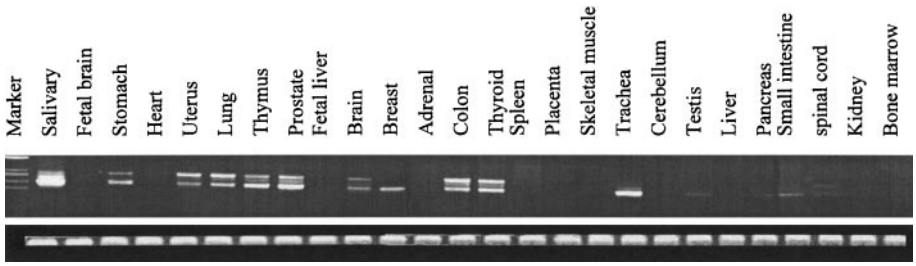

FIG. 7. Tissue expression of the KLK12 gene as determined by RT-PCR. β-actin, a housekeeping gene, was used as a control gene. The actin PCR product is 838 bp. The variant forms were best shown in prostate, colon, thymus, and thyroid tissues. The upper band (905 bp) is the classic form (see Fig. 1), the middle band (776 bp) is the related protein-1, and the lower band (644 bp) is the related protein-2. For splice variant discussion, see text. The primers used were L5-F2 and L5-R2, as shown in Table 1. For further discussion, see text.

conserved (I-II-I-0) (see legend of Fig. 3 for definition of intron phases). (d) These genes are clustered in the same chromosomal region, without any intervening non-kallikrein-like genes (Fig. 10). (e) The histidine residue of the catalytic triad of serine proteases is located near the end of the second coding exon; the aspartate residue is located in the middle of the third coding exon; and serine is found at the beginning of the fifth coding exon. As shown in Fig. 3, all these criteria apply to the newly identified KLK12 gene. Thus, KLK12 should be considered a new member of the kallikrein multigene family.

Serine proteases and kallikreins are synthesized as "preproenzymes" that contain an N-terminal signal peptide (prezymogen), followed by a short activation peptide and the enzymatic domain (Nelson *et al.*, 1999; Rittenhouse *et al.*, 1998). PreproPSA has 24 additional residues that constitute the preregion (signal peptide, 17 residues) and the propeptide (7 residues) (McCormack *et al.*, 1995). The signal peptide directs the protein to and across the endoplasmic reticulum (ER). In the ER, the signal peptide is removed, and the resulting proPSA is transported to the plasma membrane, where it is secreted. Our hydrophobicity study (Fig. 4) indicates that the amino-terminal region of the KLK12 protein harbors a signal peptide. Also, computer analysis of the amino acid sequence of KLK12 predicted a cleavage site between amino acids 17 and 18 (SQA-AT). Thus, although the protein product has not as yet been characterized, it is very likely a secreted protein. We further predicted an activation cleavage site between amino acids 21 and 22 (K $\downarrow$ I).

The DNA sequences coding for eukaryotic proteins are rarely contiguous; usually, they are separated in the genome by intervening noncoding sequences (introns). Following polyadenylation of the primary transcript, the introns are removed by splicing to generate the mature mRNA. The minimal sequences required for splicing of mRNA are located at the 5′ and 3′ boundaries of the intron. In metazoans, only the first and last two bases are highly conserved (GT . . . AG) (Sambrook *et al.*, 1989). RNA splicing provides a mechanism whereby the expression of particular proteins with specialized functions can be restricted to certain cell or tissue types during development (Adams *et al.*, 1996). Furthermore, it is now known that about 15% of mutations in mammalian genes that are implicated in disease states affect RNA splicing signals (Horowitz and Krainer, 1994). Recent literature suggests that distinct molecular forms of PSA could be expressed differently by malignant versus benign prostate epithelium (Baffa *et al.*, 1996). Aberrant PSA RNA splicing in benign prostatic hyperplasia as opposed to prostate cancer has been described by Henttu *et al.* (1990). In addition, it has been postulated that different prostatic tissues potentially harboring occult cancer could

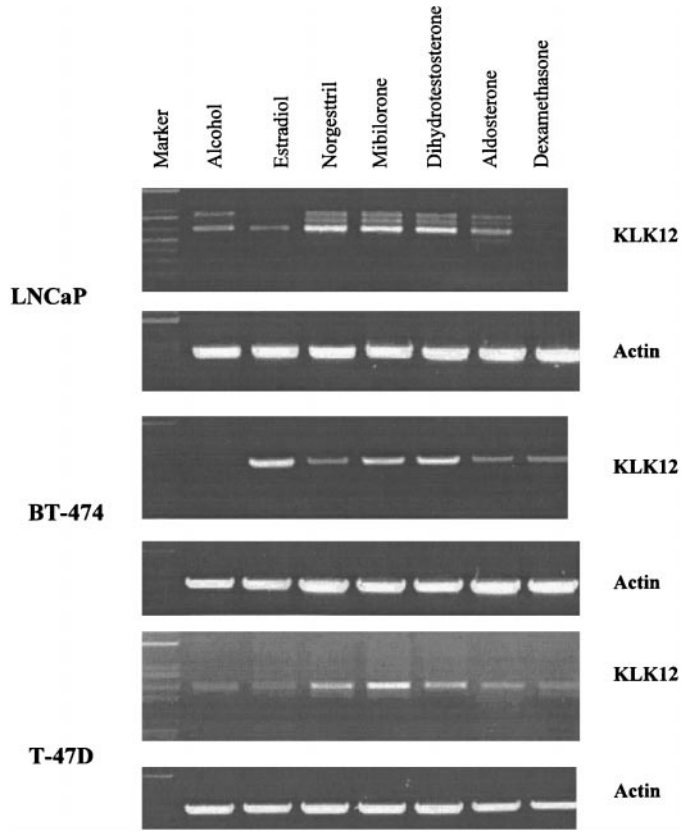

FIG. 8. Hormonal regulation of the KLK12 gene in the LNCaP prostatic carcinoma cell line and BT-474 and T-47D breast carcinoma cell lines. Steroids were at 10^{-8} M final concentration. β-actin (a housekeeping gene that is not regulated by steroid hormones) was used as a control gene. The pS2 gene (up-regulated by estrogens) and PSA (up-regulated by androgens and progestins) were used as control genes and performed as expected (data not shown). Note detection of the three isoforms only in the LNCaP cell line. For interpretations, see text.

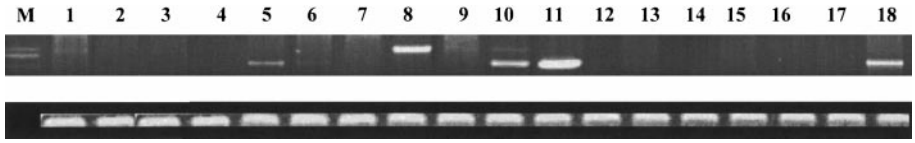

FIG. 9. Expression of the KLK12 gene in breast cancer (**1–17**) and normal (**18**) tissues. Note complete absence of expression in 12 cancer tissues. For isoforms, see also Fig. 7, and for more comments, see text. Actin was used as a contol gene.

account for the presence of various forms of PSA (Baffa *et al.*, 1996).

The presence of alternatively spliced forms is not a unique feature of the KLK12 gene; several other kallikreins are known to be expressed in various alternatively spliced forms. In addition to the major 1.6-kb transcript, several RNA species are transcribed from the PSA gene (Heuze *et al.*, 1999). Furthermore, others (Lundwall and Lilja, 1987; Riegman *et al.*, 1988) have described several PSA isoforms. In addition, Riegman *et al.* (1991) reported the identification of two alternatively spliced forms of the human glandular kallikrein (KLK2) gene, and Liu *et al.* (1999) isolated three alternative forms of the same gene. A novel transcript of the tissue kallikrein gene was isolated from the colon (Chen *et al.*, 1994). Neuropsin, a recently identified kallkrein-like gene, was found to have two alternatively spliced forms in addition to the major form (Mitsui *et al.*, 1999; Yoshida *et al.*, 1998). Here, we describe the cloning of the classical kallikrein form and two unique splice forms of the KLK12 gene. Because the classical form and the splice forms all have the same 5′ sequence required for translation, secretion, and activation as other kallikreins, it is reasonable to assume that all three mRNA forms should produce a secreted protein. At the mRNA level, it would be interesting to determine which form is predominant in a certain tissue or a certain clinical condition. Our preliminary findings have been presented in Figs. 7 and 9. It should be realized also that the truncated forms of the protein lack some of the amino acids essential for the catalytic activity of serine proteases, thus, it is unlikely that these forms will encode a secreted protein.

Kallikreins are involved in many physiological processes and also in the pathogenesis of human diseases, depending on the tissue of their primary expression. The KLK1 gene is involved in many disease processes, including inflammation (Clements, 1997), hypertension (Margolius *et al.*, 1974), renal nephritis, and diabetic renal disease (Cumming *et al.*, 1994; Jaffa *et al.*, 1992). The connections of human stratum corneum chymotryptic enzyme (HSCCE) (KLK7) with skin diseases, including pathological keratinization and psoriasis, have already been reported (Ekholm and Egelrud, 1999; Sondell *et al.*, 1996). Little *et al.* (1997) suggested that zyme (KLK6) may be amyloidogenic and may play a role in the development of Alzheimer's disease. There are other reports describing connection of neuropsin (KLK8) expression with diseases of the central nervous system, including epilepsy (Kishi *et al.*, 1999; Momota *et al.*, 1998).

Increasing evidence indicates that many of the kallikreins and kallikrein-like genes that are clustered in the same region of chromosome 19 (Fig. 10) are related to malignancy. PSA is the best marker for prostate cancer so far (Diamandis, 1998). A recent report provided evidence that PSA has antiangiogenic activity and that this activity may be related to its action as a serine protease (Fortier *et al.*, 1999). This study suggested also that other serine proteases and members of the kallikrein multigene family of enzymes should be evaluated for potential antiangiogenic action. Recent literature suggests that hK2 (encoded by the KLK2 gene) could be a useful differential diagnostic marker for certain subtypes of prostate cancer (Magklara *et al.*, 1999; Partin *et al.*, 1999; Rittenhouse *et al.*, 1998; Stenman, 1999). An *in vitro* study suggests that NES1 is a tumor suppressor gene (Goyal *et al.*, 1998). Protease M (zyme) is differentially expressed in primary breast and ovarian tumors (Anisowicz *et al.*, 1996), and the

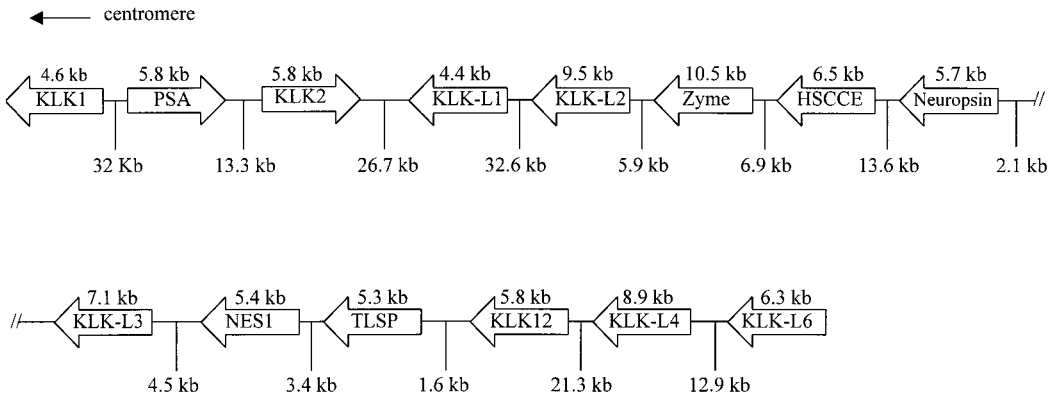

FIG. 10. An approximately 300-kb region of almost contiguous genomic sequence around chromosome 19q13.3–q13.4. Genes are represented by horizontal arrows denoting the direction of the coding sequence. Distances between genes are in kilobases. Figure is not drawn to scale. For full gene names, see abbreviations footnote.

HSCCE has been shown to be expressed at abnormally high levels in ovarian cancer (Tanimoto *et al.*, 1999). Another recently identified kallikrein-like gene, located close to KLK12, the tumor-associated differentially expressed gene-14 (TADG-14/neuropsin) was found to be overexpressed in a subset of ovarian cancer tissues (Underwood *et al.*, 1999). Furthermore, prostase/KLK-L1 is speculated to be linked to prostate cancer (Nelson *et al.*, 1999).

Our preliminary results indicate that KLK12 is upregulated by steroid hormones in breast and prostate cancer cell lines (Fig. 8). These results are not surprising, since many other kallikrein genes are also regulated by steroid hormones (for review see Diamandis *et al.*, 2000a). The differences in the rank of potency of steroid hormones among different cell lines could be attributed to differences in the abundance of the steroid hormone receptors between them as described elsewhere (Magklara *et al.*, 2000). However, these results should be considered preliminary and need further evaluation by more quantitative methods.

In conclusion, we characterized a new member of the human kallikrein gene family, KLK12, which maps to the human kallikrein locus (chromosome 19q13.3–q13.4). This gene has two related splice forms in addition to the main form. KLK12 is expressed in a variety of tissues and appears to be down-regulated in breast cancer, and its expression is influenced by steroid hormones. Since a few other kallikreins are already being used as valuable tumor markers, we speculate that KLK12 may also find a similar clinical application. This possibility, as well as the physiological function of the protein, needs further investigation.

REFERENCES

Adams, M. D., Rudner, D. Z., and Rio, D. C. (1996). Biochemistry and regulation of pre-mRNA splicing. *Curr. Opin. Cell. Biol.* **8:** 331–339.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and 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., and Sager, R. (1996). A novel protease homolog differentially expressed in breast and ovarian cancer. *Mol. Med.* **2:** 624–636.

Baffa, R., Moreno, J. G., Monne, M., Veronese, M. L., and Gomella, L. G. (1996). A comparative analysis of prostate-specific antigen gene sequence in benign and malignant prostate tissue. *Urology* **47:** 795–800.

Ballabio, A. (1993). The rise and fall of positional cloning? *Nat. Genet.* **3:** 277–279. [News]

Bhoola, K. D., Figueroa, C. D., and Worthy, K. (1992). Bioregulation of kinins: Kallikreins, kininogens, and kininases. *Pharmacol. Rev.* **44:** 1–80.

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

Clements, J. (1997). The molecular biology of the kallikreins and their roles in inflammation. *In* "The Kinin System" (S. Farmer, Ed.), pp. 71–97, Academic Press, New York.

Collins, F. S. (1995). Positional cloning moves from perditional to traditional [Published erratum appears in *Nat. Genet.* 1995 Sep; **11**(1):104]. *Nat. Genet.* **9:** 347–350.

Cumming, A. D., Walsh, T., Wojtacha, D., Fleming, S., Thomson, D., and Jenkins, D. A. (1994). Expression of tissue kallikrein in human kidney. *Clin. Sci.* **87:** 5–11.

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., and Obiezu, C. V. (2000a). The new human kallikrein gene family: Implications in carcinogenesis. *Trends Endocrinol. Metab.* **11:** 54–60.

Diamandis, E. P., Yousef, G. M., Clements, J., Ashworth, L. K., Yoshida, S., Egelrud, T., *et al.* (2000b). New nomenclature for the human tissue kallikrein gene family. *Clin. Chem.*, in press.

Ekholm, E., and Egelrud, T. (1999). Stratum corneum chymotryptic enzyme in psoriasis. *Arch. Dermatol. Res.* **291:** 195–200.

Evans, B. A., Yun, Z. X., Close, J. A., Tregear, G. W., Kitamura, N., Nakanishi, S., Callen, D. F., Baker, E., Hyland, V. J., Sutherland, G. R., *et al.* (1988). Structure and chromosomal localization of the human renal kallikrein gene. *Biochemistry* **27:** 3124–3129.

Fortier, A. H., Nelson, B. J., Grella, D. K., and Holaday, J. W. (1999). Antiangiogenic activity of prostate-specific antigen. *J. Natl. Cancer Inst.* **91:** 1635–1640.

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

Henttu, P., Lukkarinen, O., and Vihko, P. (1990). Expression of the gene coding for human prostate-specific antigen and related hGK-1 in benign and malignant tumors of the human prostate. *Int. J. Cancer* **45:** 654–660.

Heuze, N., Olayat, S., Gutman, N., Zani, M. L., and Courty, Y. (1999). Molecular cloning and expression of an alternative hKLK3 transcript coding for a variant protein of prostate-specific antigen. *Cancer Res.* **59:** 2820–2824.

Horowitz, D. S., and Krainer, A. R. (1994). Mechanisms for selecting 5′ splice sites in mammalian pre-mRNA splicing. *Trends Genet.* **10:** 100–106.

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.

Irwin, D. M., Robertson, K. A., and MacGillivray, R. T. (1988). Structure and evolution of the bovine prothrombin gene. *J. Mol. Biol.* **200:** 31–45.

Jaffa, A. A., Chai, K. X., Chao, J., Chao, L., and Mayfield, R. K. (1992). Effects of diabetes and insulin on expression of kallikrein and renin genes in the kidney. *Kidney Int.* **41:** 789–795.

Kishi, T., Kato, M., Shimizu, T., Kato, K., Matsumoto, K., Yoshida, S., Shiosaka, S., and Hakoshima, T. (1999). Crystal structure of neuropsin, a hippocampal protease involved in kindling epileptogenesis. *J. Biol. Chem.* **274:** 4220–4224.

Kozak, M. (1991). An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* **115:** 887–903.

Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996). The I.M.A.G.E. Consortium: An integrated molecular analysis of genomes and their expression. *Genomics* **33:** 151–152.

Little, S. P., Dixon, E. P., Norris, F., Buckley, W., Becker, G. W., Johnson, M., Dobbins, J. R., Wyrick, T., Miller, J. R., MacKellar, W., *et al.* (1997). Zyme, a novel and potentially amyloidogenic enzyme cDNA isolated from Alzheimer's disease brain. *J. Biol. Chem.* **272:** 25135–25142.

Liu, X. F., Essand, M., Vasmatzis, G., Lee, B., and Pastan, I. (1999). Identification of three new alternate human kallikrein 2 tran-

scripts: Evidence of long transcript and alternative splicing. *Biochem. Biophys. Res. Commun.* **264:** 833–839.

Liu, X. L., Wazer, D. E., Watanabe, K., and Band, V. (1996). Identification of a novel serine protease-like gene, the expression of which is down-regulated during breast cancer progression. *Cancer Res.* **56:** 3371–3379.

Lundwall, A., and Lilja, H. (1987). Molecular cloning of human prostate specific antigen cDNA. *FEBS Lett.* **214:** 317–322.

Luo, L., Herbrick, J. A., Scherer, S. W., Beatty, B., Squire, J., and Diamandis, E. P. (1998). Structural characterization and mapping of the normal epithelial cell-specific 1 gene. *Biochem. Biophys. Res. Commun.* **247:** 580–586.

Magklara, A., Grass, L., and Diamandis, E. P. (2000). Differential steroid hormone regulation of human glandular kallikrein (hK2) and prostate specific antigen (PSA) in breast cancer cell lines. *Breast Cancer Res. Treat.* **59:** 263–270.

Magklara, A., Scorilas, A., Catalona, W. J., and Diamandis, E. P. (1999). The combination of human glandular kallikrein and free prostate-specific antigen (PSA) enhances discrimination between prostate cancer and benign prostatic hyperplasia in patients with moderately increased total PSA. *Clin. Chem.* **45:** 1960–1966.

Margolius, H. S., Horwitz, D., Pisano, J. J., and Keiser, H. R. (1974). Urinary kallikrein excretion in hypertensive man. Relationships to sodium intake and sodium-retaining steroids. *Circ. Res.* **35:** 820–825.

McCormack, R. T., Rittenhouse, H. G., Finlay, J. A., Sokoloff, R. L., Wang, T. J., Wolfert, R. L., Lilja, H., and Oesterling, J. E. (1995). Molecular forms of prostate-specific antigen and the human kallikrein gene family: A new era. *Urology* **45:** 729–744.

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

Momota, Y., Yoshida, S., Ito, J., Shibata, M., Kato, K., Sakurai, K., Matsumoto, K., and Shiosaka, S. (1998). Blockade of neuropsin, a serine protease, ameliorates kindling epilepsy. *Eur. J. Neurosci.* **10:** 760–764.

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

Partin, A. W., Catalona, W. J., Finlay, J. A., Darte, C., Tindall, D. J., Young, C. Y., Klee, G. G., Chan, D. W., Rittenhouse, H. G., Wolfert, R. L., and Woodrum, D. L. (1999). Use of human glandular kallikrein 2 for the detection of prostate cancer: Preliminary analysis. *Urology* **54:** 839–845.

Riegman, P. H., Klaassen, P., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1988). Molecular cloning and characterization of novel prostate antigen cDNA's. *Biochem. Biophys. Res. Commun.* **155:** 181–188.

Riegman, P. H., Vlietstra, R. J., van der Korput, H. A., Romijn, J. C., and 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.

Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1989). Characterization of the prostate-specific antigen gene: A novel human kallikrein-like gene. *Biochem. Biophys. Res. Commun.* **159:** 95–102.

Rittenhouse, H. G., Finlay, J. A., Mikolajczyk, S. D., and 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.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Sping Harbor Laboratory Press, Cold Spring Harbor, NY.

Schachter, M. (1979). Kallikreins (kininogenases)—A group of serine proteases with bioregulatory actions. *Pharmacol. Rev.* **31:** 1–17.

Schedlich, L. J., Bennetts, B. H., and Morris, B. J. (1987). Primary structure of a human glandular kallikrein gene. *DNA* **6:** 429–437.

Sondell, B., Dyberg, P., Anneroth, G. K., Ostman, P. O., and Egelrud, T. (1996). Association between expression of stratum corneum chymotryptic enzyme and pathological keratinization in human oral mucosa. *Acta Derm. Venereol.* **76:** 177–181.

Stenman, U. H. (1999). New ultrasensitive assays facilitate studies on the role of human glandular kallikrein (hK2) as a marker for prostatic disease. *Clin. Chem.* **45:** 753–754.

Tanimoto, H., Underwood, L. J., Shigemasa, K., Yan Yan, M. S., Clarke, J., Parmley, T. H., and 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.

Underwood, L. J., Tanimoto, H., Wang, Y., Shigemasa, K., Parmley, T. H., and O'Brien, T. J. (1999). Cloning of tumor-associated differentially expressed gene-14, a novel serine protease overexpressed by ovarian carcinoma. *Cancer Res.* **59:** 4435–4439.

Yamashiro, K., Tsuruoka, N., Kodama, S., Tsujimoto, M., Yamamura, Y., Tanaka, T., Nakazato, H., and Yamaguchi, N. (1997). Molecular cloning of a novel trypsin-like serine protease (neurosin) preferentially expressed in brain. *Biochim. Biophys. Acta* **1350:** 11–14.

Yoshida, S., Taniguchi, M., Hirata, A., and Shiosaka, S. (1998). Sequence analysis and expression of human neuropsin cDNA and gene. *Gene* **213:** 9–16.

Yousef, G. M., Chang, A., and Diamandis, E. P. (2000). Identification and characterization of KLK-L4, a new kallikrein-like gene that appears to be down-regulated in breast cancer tissues. *J. Biol. Chem.* **275**(16): 11891–11898.

Yousef, G. M., and Diamandis, E. P. (1999). The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation. *J. Biol. Chem.* **274:** 37511–37516.

Yousef, G. M., and Diamandis, E. P. (2000). The expanded human kallikrein gene family: Locus characterization and molecular cloning of a new member, KLK-L3 (KLK9). *Genomics* **65:** 184–194.

Yousef, G. M., Luo, L. Y., Scherer, S. W., Sotiropoulou, G., and Diamandis, E. P. (1999a). Molecular characterization of Zyme/Protease M/Neurosin (PRSS9), a hormonally regulated kallikrein-like serine protease. *Genomics* **62:** 251–259.

Yousef, G. M., Obiezu, C. V., Luo, L. Y., Black, M. H., and Diamandis, E. P. (1999b). Prostase/KLK-L1 is a new member of the human kallikrein gene family, is expressed in prostate and breast tissues, and is hormonally regulated. *Cancer Res.* **59:** 4252–4256.