Biol. Chem., Vol. 382, pp. 5–14, January 2001 · Copyright © by Walter de Gruyter · Berlin · New York

Review

The Expanded Human Kallikrein *(KLK)* Gene Family: Genomic Organisation, Tissue-Specific Expression and Potential Functions

Judith Clements*, John Hooper, Ying Dong and Tracey Harvey

Centre for Molecular Biotechnology, School of Life Sciences, Queensland University of Technology, Brisbane, Australia 4001

* Corresponding author

The tissue kallikreins are serine proteases encoded by highly conserved multi-gene families. The rodent kallikrein *(KLK)* families are particularly large, consisting of 13–26 genes clustered in one chromosomal locus. It has been recently recognised that the human *KLK* gene family is of a similar size (15 genes) with the identification of another 12 related genes *(KLK4-KLK15)* within and adjacent to the original human *KLK* locus *(KLK1-3)* on chromosome 19q13.4. The structural organisation and size of these new genes is similar to that of other *KLK* genes except for additional exons encoding 5' or 3' untranslated regions. Moreover, many of these genes have multiple mRNA transcripts, a trait not observed with rodent genes. Unlike all other kallikreins, the *KLK4-KLK15* encoded proteases are less related (25–44%) and do not contain a conventional kallikrein loop. Clusters of genes exhibit high prostatic *(KLK2-4, KLK15)* or pancreatic *(KLK6-13)* expression, suggesting evolutionary conservation of elements conferring tissue specificity. These genes are also expressed, to varying degrees, in a wider range of tissues suggesting a functional involvement of these newer human kallikrein proteases in a diverse range of physiological processes.

Key words: Gene family/Gene structure/*KLK*/Tissue Kallikrein/Tissue specificity.

Introduction

The tissue kallikreins are a subgroup of serine proteasesw that are characterised by their homology to true tissue kallikrein, encoded by the *KLK1* gene (Clements, 1989, 1997; Bhoola *et al.,* 1992). Typically, they share a high degree of sequence and structural similarity and, at least in three species – mouse, rat and man – the genes are clustered together at one locus (Evans *et al.,* 1987; Southard-Smith *et al.,* 1994; Kurtz and St.Lezin, 1992; Harvey *et al.,* 2000; Yousef *et al.,* 2000a). Although highly conserved structurally, their known or predicted enzymic functions are specific and distinct from one another. They also share a wide range of expression patterns that likely indicate involvement in a diverse range of physiological processes (Bhoola *et al.,* 1992; Clements, 1997; Rittenhouse *et al.,* 1998; Diamandis *et al.,* 2000).

In contrast to the rodent tissue kallikrein *(KLK)* gene families, the human *KLK* gene family appeared to consist of just three genes – *KLK1* encoding tissue kallikrein, *KLK2* encoding glandular kallikrein and *KLK3*, which encodes prostate-specific antigen or PSA (Riegman *et al.,* 1992; Clements, 1994). These enzymes play important roles in the (patho)physiology of the kidney, pancreas and brain and cardiovascular, respiratory, gastrointestinal and reproductive systems (tissue kallikrein), prostate (PSA and K2) and breast (PSA) (Bhoola *et al.,* 1992; Clements, 1997; Rittenhouse *et al.,* 1998; Black and Diamandis, 2000). Recently, several groups have identified other human serine proteases with homology to tissue kallikrein or PSA (Dihanich and Speiss, 1994; Hansson *et al.,* 1994; Anisowicz *et al.,* 1996; Liu *et al.,* 1996; Little *et al.,* 1997; Yamashiro *et al.,* 1997; Yoshida *et al.,* 1998a,b; Brattsand and Egelrud, 1999; Nelson *et al.,* 1999; Stephenson *et al.,* 1999; Tanimoto *et al.,* 1999; Underwood *et al.,* 1999; Hu *et al.,* 2000; Mitsui *et al.,* 2000; Yousef and Diamandis 1999, 2000; Yousef *et al.,* 1999a-c; 2000b-e). All these genes were clustered at the human *KLK* locus on chromosome 19q13.4, thus expanding the human *KLK* gene family to at least 15 genes (Stephenson *et al.,* 1999; Gan *et al.,* 2000; Harvey *et al.,* 2000; Yousef *et al.,* 1999a, 2000a). In this review, we summarise briefly the current knowledge on the structure and function of these genes, their genomic organisation, sites of expression and known or predicted physiological roles.

The Human *KLK* Locus at Proximal 19q13.4

The first three characterised human *KLK* genes – *KLK1, KLK2* and *KLK3* – were localised in a cluster to a 60 kilobase region at q13.3–13.4 on chromosome 19 (Riegman *et al.,* 1992; Stephenson *et al.,* 1999). This region is syntenic to the locus on mouse chromosome 7 where the mouse *KLK* gene family cluster is localised, further emphasising evolutionary similarities between these families (Howles *et al.,* 1984; Evans *et al.,* 1987). However, the

striking difference between the mouse (and rat) and human *KLK* families was their size. The rodent families are large consisting of 13–26 genes (Evans *et al.,* 1987; Wines *et al.,* 1989); the rat family is also clustered in one locus on rat chromosome 1 (Kurtz and St.Lezin, 1992; Southard-Smith *et al.,* 1994). Thus, by comparison, it appeared logical that the human *KLK* family may also contain more genes specifically within, or adjacent to, the original *KLK* cluster on 19q13.3–13.4. Early genomic Southern blot analyses presented conflicting results with most groups suggesting there were only 3–4 genes (Baker and Shine, 1985; Fukushima *et al.,* 1985; Evans *et al.,* 1988; Riegman *et al.,* 1989). However, Southern blot analyses of a human genomic library, using a monkey *KLK* cDNA probe with 92–95% homology to human *KLK* cDNAs, indicated that there may be as many as 19 human *KLK* genes (Murray *et al.,* 1990).

Subsequently, several newly characterised human serine proteases – Protease M/neurosin/zyme, normal epithelial cell-specific 1 (Nes 1), trypsin-like serine protease (TLSP), stratum corneum chymotryptic enzyme (SCCE), and neuropsin – were identified to have some homology (~40%) to tissue kallikrein or PSA (Hansson *et al.,* 1994; Anisowicz *et al.,* 1996; Liu *et al.,* 1996; Little *et al.,* 1997; Yamashiro *et al.,* 1997; Yoshida *et al.,* 1998a,b). Although these enzymes are less related than tissue kallikrein, K2 and PSA (62–77%) are to each other, two of the encoding genes were localised by FISH analysis to 19q13.3–13.4, in the vicinity of the *KLK* locus (Anisowicz *et al.,* 1996; Little *et al.,* 1997; Luo *et al.,* 1998). This suggested that indeed there were other serine protease genes in this region but that they were less related to the *KLK1-3* encoded proteases in sharp comparison to the high conservation observed between the rodent kallikreins (70–85%) (Evans *et al.,* 1987; Wines *et al.,* 1989).

Recently, both the availability of the draft sequence for this region from the Human Genome Sequencing Project (at Lawrence Livermore National Laboratory, California, USA) and physical mapping has allowed the identification of 12 additional kallikrein-related genes *(KLK4–15)* within a 320 kilobase region that encompasses the original *KLK* locus (Gan *et al.,* 2000; Harvey *et al.,* 2000; Yousef *et al.,* 1999a, 2000a). Although many of these new genes/proteases are known by a variety of designations (see above and Table 1), the Human Genome Nomenclature Committee recently approved the use of the consistent *KLK* nomenclature to describe these genes (Diamandis *et al.,* 2000b). *KLK4–14* lie telomeric of the original locus and are transcribed telomere to centromere in the same direction as *KLK1* (Figure 1). *KLK15* lies between *KLK1* and *KLK3* and is also transcribed in the same direction as *KLK1. KLK2* and *KLK3* are the only two genes in the locus that are transcribed in the opposite direction – centromere to telomere. Like the rodent families that contain 3 (rat) to 10 (mouse) pseudogenes (Berg *et al.,* 1992), 5 pseudogene sequences have been identified surrounding the *KLK3-KLK2-KLK4* cluster at the cen-

Table 1 Summary of the mRNA Expression Patterns of the 15 *KLK* Genes Obtained by Northern Blot and RT-PCR Analysis.

		Northern and mRNA dot blot tissue expression		RT-PCR tissue expression	
Current name	Other names	High	Moderate	High	Moderate
KLK1	Tissue kallikrein	Kidney Pancreas Salivary gland		Endometrium ◆ Kidney ◆ Ovary ◆	Prostate ◆ Salivary gland ◆
KLK2	Glandular kallikrein	Prostate		Prostate ◆	
KLK3	PSA	Prostate		Prostate ◆	
KLK4	Prostase PRSS17 KLK-like1	Prostate		Prostate ◆ Ovary ◆	Endometrium ◆ Salivary gland ◆ Kidney ◆ Uterus Breast Adrenal gland Colon Thyroid Testis
KLK5	KLK-like SCTE	Testis Breast Skin	Salivary gland Esophagus	Skin ◆ Endometrium ◆ Ovary ◆	Breast ◆ Kidney Brain
KLK6	Protease M Neurosin Zyme PRSS9	Pancreas Brain Kidney	Spinal cord Testis Appendix Colorectal adenocarcinoma	Breast ◆ Skin ◆ Endometrium ◆ Kidney Ovary ◆	Salivary gland Heart Thyroid Spleen Placenta

Table 1 (continued)

		Northern and mRNA dot blot tissue expression		RT-PCR tissue expression	
Current name	Other names	High	Moderate	High	Moderate
				Prostate ◆ Brain Uterus Thymus Spinal cord Kidney	Testis Trachea
KLK7	SCCE PRSS6	Pancreas Skin		Breast ◆ Skin ◆ Endometrium ◆ Ovary ◆	Prostate Kidney
KLK8	Neuropsin TADG-14 Ovasin PRSS19	Pancreas Brain Skin	Duodenum Liver Esophagus Stomach Salivary gland	Endometrium ◆ Ovary ◆ Skin ◆	Breast ◆ Prostate ◆ Salivary gland
KLK9	KLK-like 3	Pancreas		Salivary gland Thymus Testis Spinal cord	Prostate Breast Trachea
KLK10	Nes 1 PRSSL1	Pancreas Breast Ovary Colon Small intestine Testis Lung	Esophagus Duodenum Trachea Colorectal adenocarcinoma Prostate	Breast ◆ Skin ◆ Endometrium ◆ Prostate Ovary ◆	
KLK11	TLSP PRSS20 Hippostasin	Pancreas Skin Prostate Lung Heart Testis Brain	Salivary gland Stomach Liver Skeletal muscle	Skin ◆ Prostate Salivary gland Ovary ◆ Brain	Breast ◆ Stomach Uterus Lung Spleen Cerebellum
KLK12	KLK-like 5	Pancreas	Brain Duodenum Appendix	Salivary gland ◆	
KLK13	KLK-like 4	Pancreas Esophagus Appendix	Duodenum Stomach Brain Testis Prostate Salivary gland	Endometrium ◆ Prostate	Breast ◆ Skin ◆ Salivary gland Ovary ◆
KLK14	KLK-like 6	Brain Bone marrow Fetal liver	Liver Pancreas Fetal spleen Prostate	Breast ◆ Skin ◆ Endometrium ◆ Prostate Salivary gland ◆ Ovary ◆	
KLK15	*KLK15*			Salivary gland Prostate Thyroid	Adrenal gland Colon Testis Kidney

The Northern blot expression patterns are mostly from tissues except for the cell line, colorectal adenocarcinoma, whereas the RT-PCR patterns are from tissues and carcinoma cell lines; (◆) represents carcinoma cell lines. All references used in the preparation of this Table are highlighted throughout the text.

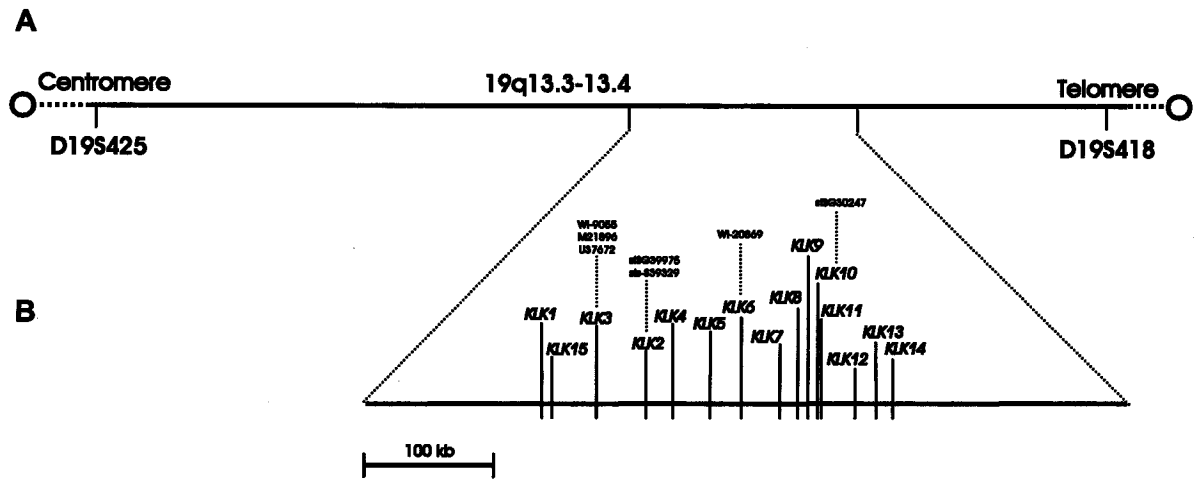

Fig. 1 Location of the *KLK* Locus at Chromosome 19q13.3–13.4.
(A) Schematic representation of the interval between D19S425 and D19S418 including 19q13.3–13.4. The *KLK* locus is located proximal to D19S418. B. The position of the 15 kallikrein encoding genes on the *KLK* locus are marked. *KLK1* and *KLK4* to *KLK15* are transcribed telomere to centromere, whereas *KLK2* and *KLK3* are transcribed in the opposite direction. The position of microsatellite markers relative to these genes in the *KLK* locus are also indicated.

tromeric end of the locus (Stephenson *et al.,* 1999; Gan *et al.,* 2000). The locus lies between intervals D19S425 to D19S418 at proximal q13.4. Specific microsatellite markers that are associated with the locus are WI-9055, M21896 and U37672 *(KLK3)*; stSG39975 and sts-S39329 *(KLK2)*; WI-20869 *(KLK6)* and stSG30247 *(KLK10)* (Harvey *et al.,* 2000) (Figure 1).

Genomic Southern blot analysis of bacterial artificial chromosomes (BACs) containing DNA from this region, with a degenerate probe to the histidine-encoding region of serine proteases, suggests that most, if not all, serine protease genes in this region have been identified (Harvey *et al.,* 2000). In support of this, the next gene telomeric of *KLK14* is an unrelated gene, SIGLEC9, which encodes a sialic acid-binding Ig-like lectin that is also a member of the immunoglobulin superfamily (Foussias *et al.,* 2000). However, draft sequence is not yet available for the region centromeric of *KLK1* and *KLK3*, thus there may still be further as yet unidentified genes within the *KLK* locus.

Structural and Sequence Similarities

The size of the human *KLK* genes range from 4–10 kilobases with most of the differences relating to differences in intron sizes, although the intron phases are completely conserved (Luo *et al.,* 1998; Nelson *et al.,* 1999; Stephenson *et al.,* 1999, Yousef and Diamandis 1999, 2000; Yousef *et al.,* 1999b,c, 2000a–e; Hu *et al.,* 2000) (Table 2). Of note, the five coding exons for these serine protease genes are highly conserved both in size and organisation. The first of these has a short 5' untranslated region prior to the initiation of the pre-signal peptide coding region. The second exon encodes the remainder of the pre- as well as pro-enzyme sequences. The mature enzyme is encoded on the remainder of the second coding exon through to the fifth coding exon with a variable sized 3' untranslated region completing this last coding exon. The positions, in the second, third and fifth coding exons, of the three residues of the catalytic triad (histidine, aspartate, serine) critical for protease activity are highly conserved across all 15 genes. Interestingly, this structural organisation is also highly conserved for all rodent *KLK* genes and the dog *KLK2* gene (Evans *et al.,* 1987, Wines *et al.,* 1989, 1991; Chapdelaine *et al.,* 1991).

The major structural difference between these newer human *KLK* genes and the original *KLK1–3* genes or those of other species is the presence of additional exons either 5' or 3' of the coding exons for many of these genes (Luo *et al.,* 1998; Hu *et al.,* 2000; Mitsui *et al.,* 2000; Yousef and Diamandis 1999; Yousef *et al.,* 1999b; 200a,c,d,e). Since the transcriptional start site or TATA box elements have not yet been delineated for all of these genes, it is possible that additional 5' exons may be present in more genes. The mRNA transcript sizes vary from ~870 to ~1600 nucleotides with variable 5' and 3' untranslated regions determining the variable length. Both classical (AATAAA) and variant (AGTAAA, ACTAAA, TATAAA, ATTAAA) polyadenylation signals are utilised.

Multiple variant mRNA transcripts, the result of alternative splicing, have also been described for many of the human *KLK* genes (Chen *et al.,* 1994; Rae *et al.,* 1999; Riegman *et al.,* 1988, 1991; Liu *et al.,* 1999; Heuze *et al.,* 1999; Tanaka *et al.,* 2000; Yousef *et al.,* 2000b,d,e; Mitsui *et al.,* 2000). This is in sharp contrast to the rodent families where only one mRNA transcript has been described for each gene (Evans *et al.,* 1987; Wines *et al.,* 1989). These splicing events cause a shift in the open reading frame, leading to premature stop codons and truncated proteins

Table 2 Comparison of the Genomic Structures of the Human Kallikrein Genes *KLK1–KLK15*.

Gene	**Exon**	Intron	**Exon**	Intron	**[5'UTR] Exon**	Intron	**His Exon**	Intron	**Asp Exon**	Intron	**Exon**	Intron	**Ser Exon [3'UTR]**	Intron	Exon
KLK1					**[39] 46**	1829 (I)	**160**	1269 (II)	**290**	118 (I)	**137**	548 (O)	**156 [46]**		
KLK2					**[29] 46**	1206 (I)	**160**	1589 (II)	**287**	113 (I)	**137**	1393 (O)	**156 [674]**		
KLK3					**[41] 46**	1238 (I)	**160**	1637 (II)	**287**	143 (I)	**137**	1376 (O)	**156 [637]**		
KLK4			49	587	**[11] 61**	1263 (I)	**163**	421 (II)	**251**	83 (I)	**137**	1272 (O)	**153 [603]**		
KLK5			31	78	**[11] 73**	2468 (I)	**262**	739 (II)	**257**	85 (I)	**134**	4853 (O)	**156 [323]**		
KLK6	185	898	52	423	**[8] 40**	739 (I)	**157**	3519 (II)	**248**	1421 (I)	**137**	1877 (O)	**153 [526]**		
KLK7			30	1393	**[58] 73**	412 (I)	**148**	696 (II)	**248**	315 (I)	**137**	2096 (O)	**156 [748]**		
KLK8			171	356	**[8] 70**	514 (I)	**160**	165 (II)	**263**	2097 (I)	**134**	1536 (O)	**156 [45]**		
KLK9					**[87] 43**	165 (I)	**157**	2459 (II)	**160**	2723 (I)	**137**	443 (O)	**150 [598]**		
KLK10			73	484	**[9] 88**	1753 (I)	**181**	953 (II)	**275**	331 (I)	**134**	464 (O)	**153 [529]**		
KLK11			86	1790	**[35] 40**	797 (I)	**157**	327 (II)	**266**	816 (I)	**137**	394 (O)	**153 [312]**		
KLK12			98	154	**[19] 37**	445 (I)	**160**	1844 (II)	**260**	954 (I)	**134**	1330 (O)	**156 [210]**		
KLK13					**[43] 52**	4396 (I)	**187**	339 (II)	**269**	1150 (I)	**137**	1762 (O)	**189 [381]**		
KLK14[a]			69	85	**[22] 40**	871 (I)	**172**	1829 (II)	**254**	497 (I)	**137**	655 (O)	**153 [177]**	254	259
KLK15					**[55] 43**	3635 (I)	**154**	500 (II)	**284**	120 (I)	**137**	673 (O)	**153 [507]**		

Bracketed numeral indicates intron phase. Coding exons are indicated by bold letters.
[a] Unpublished data: KLK14 Genbank accession number AF283669.

if translated. These variant transcripts would not encode functional serine proteases as typically one or other of the three residues critical for catalytic activity is missing. The degree of tissue-specificity in expression of these alternatively spliced mRNAs or protein products or their relationship to a particular pathophysiological condition is yet to be determined. Whether these events represent a means of further functional diversification of the human *KLK* gene family or is a control mechanism for regulating levels of the native protein or is simply an unrelated genomic processing event requires further clarification.

In contrast to the rodent *KLK/Klk* families and the human *KLK1–KLK3* encoded enzymes which are highly conserved (rodents: 70–85%; K1–K2: 62–77%), the newer human K4–K15 enzymes are less related (25–44%) (Figure 2). This may reflect an evolutionary order of gene duplication within this locus with the *KLK1–KLK3* genes perhaps the result of the most recent duplication. In keeping with this observation, phylogenetic analysis of the human *KLK* locus suggests that there are a number of sub-groups within this family (Figure 3). As noted above, key residues (and much of the adjacent regions) that denote serine protease activity are entirely conserved (Figure 2). The majority of these enzymes are trypsin-like in action, a function denoted by an aspartate six residues before the catalytic serine; four enzymes (K3, K7, K9, K15) have alternate residues in this position which likely denotes a chymotrypsin-like action. Ten of the disulfide bond-forming cysteines are absolutely conserved amongst the 15 kallikreins. A sixth bridge forming cysteine pair is conserved in K4–K12 and K14–K15, but is lacking in K1–K3 and K13 (Figure 2). K1–K3 also contain an insertion of 11 amino acids prior to the catalytic aspartate, denoted the kallikrein loop, that is thought to be important for substrate specificity of these enzymes. The newer proteases do not have this loop region although five of them have smaller insertions (Figure 2) that also may be important in specifying their enzymic action.

Tissue-Specific Patterns of Expression

As with the rat *KLK* family (MacDonald *et al.,* 1996), the newer members of the human *KLK* gene family are expressed in a large number of tissues, particularly when expression patterns are examined using the sensitive reverse-transcription-polymerase chain reaction (RT-PCR) (Brattsand and Egelrud, 1999; Yousef and Diamandis 1999, 2000; Yousef *et al.,* 1999b,c, 2000b–e; Harvey *et al.,* 2000) (Table 1). However, the number of tissues in which they are most highly expressed, as delineated using Northern and mRNA dot blot analysis, is more restricted and surprisingly conserved. In comparison to the well described *KLK1* gene, that is highly expressed in the salivary gland, kidney and pancreas (Baker and Shine, 1985; Fukushima *et al.,* 1985; Evans *et al.,* 1988), no other genes exhibit that precise pattern of expression, although, the *KLK6* gene is highly expressed in the brain, kidney and pancreas (Anisowicz *et al.,* 1996; Little *et al.,* 1997; Harvey *et al.,* 2000). Additionally, several other genes are highly expressed in the pancreas *(KLK7–13*; Liu *et al.,* 1996; Harvey *et al.,* 2000). Of interest is the clustering of these eight genes within the locus perhaps indicating conservation of promoter elements conferring tissue specific expression or a common locus control region. Indeed, a pancreatic-specific-like sequence similar to that identified in other rat pancreatic serine proteases (Swift *et al.,* 1984; Boulet *et al.,* 1986) has been identified in the promoter of the rodent and human *KLK1* genes (Wines *et al.,* 1989).

Fig. 2 Multiple Sequence Alignment of Human Kallikrein Serine Proteases.
Residues which are identical between at least eight of the enzymes are boxed. Catalytic residues are indicated by filled boxes. Disulfide bond forming cysteines are indicated by filled circles. The residue located at the bottom of the substrate binding pocket is indicated by a circle. The kallikrein loop region is overlined. The activation site is indicated by a filled triangle. The alignment was performed using the PileUp program. GenBank accession numbers: K1, M33105; K2, M18157; K3, M26663; K4, AF148532; K5, AF135028; K6, U62801; K7, L33404; K8, AB009849; K9, AF135026; K10, AF024605; K11, AB012917; K12, AF135025; K13, AF135024; K14, AF283669; K15, AF242195.

Furthermore, the *KLK4* gene exhibits an identical expression pattern to that of the previously characterised *KLK2* and *KLK3* genes with abundant expression restricted to the prostate (Chapdelaine *et al.,* 1988; Riegman *et al.,* 1988; Nelson *et al.,* 1999; Stephenson *et al.,* 1999, Harvey *et al.,* 2000). *KLK2–4* are clustered together in the locus and the promoters of *KLK2* and *KLK3* have been extensively characterised and show a high degree of conservation (Wolf *et al.,* 1992; Murtha *et al.,* 1993; Cleutjens *et al.,* 1997; Sun *et al.,* 1997). By Northern and mRNA dot blot analysis, other tissues in which *KLK* expression is abundant are the brain *(KLK8, KLK11, KLK14)*, skin *(KLK5, KLK7, KLK8, KLK11)*, lung *(KLK10, KLK11),* heart *(KLK11)*, testis *(KLK5, KLK10, KLK11),* mammary gland *(KLK5, KLK10),* ovary, small intestine and colon *(KLK10)* (Anisowicz *et al.,* 1996; Liu *et al.,* 1996; Yoshida *et al.,* 1998; Brattsand and Egelrud, 1999; Harvey *et al.,* 2000; Mitsui *et al.,* 2000). *KLK5, KLK6, KLK8, KLK10* are also expressed at high levels, at the mRNA level, in ovarian or breast cancer carcinomas or cell lines (Anisowicz *et al.,* 1996; Liu *et al.,* 1996; Tanimoto *et al.,* 1999; Underwood *et al.,* 1999). In addition to these sites

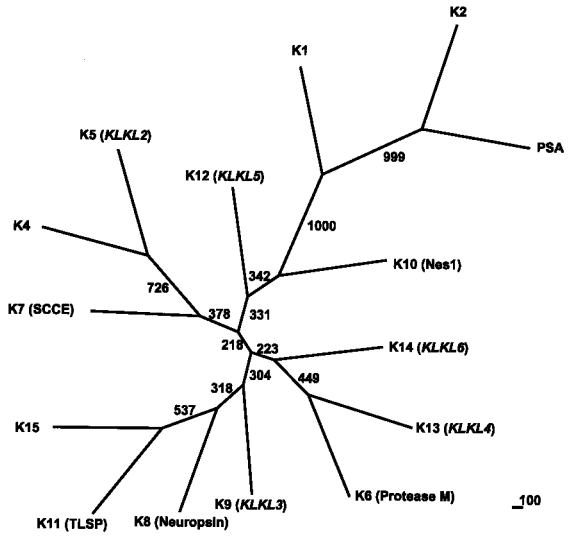

Fig. 3 Phylogenetic Tree of the Expanded KLK Family.
The tree was generated from an Eclustal multiple sequence alignment of the 15 kallikrein proteins using the Dayhoff program of the Phylip package. The lower the branch number the more closely related the genes.

of high expression, RT-PCR analysis demonstrates that these newer *KLK* genes are also expressed in a range of other tissues including salivary gland, trachea, esophagus, spinal cord, thymus, appendix, spleen, bone marrow, adrenal gland, thyroid, uterus, placenta and fetal organs as well as several carcinoma cell lines (Brattsund and Egelrud, 1999; Yousef and Diamandis, 1999, 2000; Yousef *et al.,* 1999a–c; 2000b–e; Harvey *et al.,* 2000).

Known and Putative Functions

From the combined Northern blot and RT-PCR data, it is clear that this family of genes is expressed to varying degrees in a large number of tissues and are likely to be important in a diverse range of patho(physiological) processes. The roles of the first three *KLK* family members to be identified – tissue kallikrein or K1, K2 and PSA – are well described. Tissue kallikrein, through the generation of (lys)bradykinin, and a ubiquitous expression pattern, is involved in the function of most systems in the body (Bhoola *et al.,* 1992; Clements, 1997). There is some *in vitro* evidence that tissue kallikrein may act on a range of other substrates such as the matrix metalloproteases (Tschesche *et al.,* 1989), but these are not proven biological functions. PSA hydrolyses seminal vesicle proteins in ejaculate, an event that is integral to sperm motility (Lilja, 1985). Like tissue kallikrein, PSA is also suggested to act on several other substrates (Rittenhouse *et al.,* 1998) but the biological relevance of these findings is not yet clear. PSA is widely used as a diagnostic marker for prostate cancer and monitoring tumour recurrence (Frydenberg *et al.,* 1997). Of interest, PSA expression is also associated with breast disease (Yu *et al.,* 1996) and PSA has been recently suggested to be a useful prognostic marker for breast cancer (Black and Diamandis, 2000). K2, to a lesser degree, is also being examined as a diagnostic/prognostic marker for prostate disease (Young *et al.,* Recker *et al.,* 1998). K2 activates pro-urinary plasminogen activator and thus plays a role in matrix degradation and cancer invasion (Frenette *et al.,* 1997). K2 can also activate pro-PSA (Lovgren *et al.,* 1997; Kumar *et al.,* 1997; Takayama *et al.,* 1997), demonstrating the close relationship of these two prostatic kallikreins.

The enzymatic functions of nearly all of the newer *KLK*-encoded proteins are not yet known. The exception is K6 or zyme that can hydrolyse amyloid precursor protein and thus is hypothesised to be important in the deposition of amyloid plaques in Alzheimer's disease (Little *et al.,* 1997). Interestingly, the *KLK6* gene is also highly expressed in breast and ovarian cancers (Anisowicz *et al.,* 1996; Yousef *et al.,* 1999b). Although the function of K6 in these malignancies is yet to be established, these findings highlight the potential for diverse, perhaps tissue-specific, roles for these enzymes. As noted above, the expression of several other *KLK* genes in breast, ovarian and prostate cancer tissues or cell lines (Anisowicz *et al.,* 1996; Nelson *et al.,* 1999; Tanimoto *et al.,* 1999; Underwood *et al.,* 1999; Yousef and Diamandis, 1999, 2000; Yousef *et al.,* 1999b, c; 2000b–e; Gan *et al.,* 2000; Harvey *et al.,* 2000) demonstrates that many members of this family of serine proteases are associated with hormone-dependent cancers (Diamandis *et al.,* 2000). Whether these enzymes, like PSA and K2, may be useful diagnostic or prognostic markers for these diseases or play important roles in cancer progression is yet to be established. Interestingly, *KLK10* is suggested to play a tumour suppressor role in breast and prostate cancer (Goyal *et al.,* 1998).

Similarly, the isolation or cloning of these enzymes from specific tissues such as the skin *(KLK5, KLK7, KLK11)* and brain *(KLK6, KLK8, KLK11)* suggests a role at these sites, although this is yet to be elucidated. Several of these genes *(KLK5, KLK6, KLK9)* are also expressed in psoriatic lesions further emphasising their involvement in the (patho)physiology of the skin (Gan *et al.,* 2000). Of interest, in a manner analagous to that of the prostatic K2 and PSA enzymes, the co-localisation of the trypsin-like K5 and chymotrypsin-like K7 enzymes in skin has led to the proposal that K5 may be the activator of K7 in this tissue (Brattsand and Egelrud, 1999). As noted above, the high expression of the *KLK6–KLK13* cluster in the pancreas (Harvey *et al.,* 2000) presumably also suggests an important role in pancreatic function for these enzymes. Other functions, in teeth development and neural plasticity, have been suggested for the K4 and K8 enzymes respectively, by extrapolation of the known actions of their putative mouse orthologues – enamel matrix serine protease and neuropsin (Hu *et al.,* 2000; Yoshida *et al.,* 1998).

Conclusion

The expanded human *KLK* serine proteases, although encoded by a large multigene family as in rodents, are less conserved than their rodent counterparts. Their largely disparate but overlapping expression patterns indicate that these enzymes are likely to be involved in the (patho)physiology of a range of different organs. The challenge now is to precisely determine the enzymic action of these proteases and their biological roles in these tissues and to assess their viability as diagnostic markers for a number of disease states.

Acknowledgements

The authors are supported by the National Health and Medical Research Council of Australia, Queensland Cancer Fund and Queensland University of Technology.

References

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.

Baker, A.R., and Shine, J. (1985). Human kidney kallikrein: cDNA cloning and sequence analysis. DNA *4*, 445–450.

Berg, T., Bradshaw, R.A., Carretero, O.A., Chao, J., Chao, L., Clements, J.A., Fahnestock, M., Fritz, H., Gauthier, F., MacDonald, R.J., Margolius, H.S., Morris, B.J., Richards, R.I., and Scicli, A.G. (1992). A common nomenclature for members of the tissue (glandular) kallikrein gene families. Agents Actions *38*, 19–25.

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

Black, M.H., and Diamandis, E.P. (2000). The diagnostic and prognostic utility of prostate-specific antigen for diseases of the breast. Breast Cancer Res. *59*, 1–14.

Boulet, A.M., Erwin, C.R., and Rutter, W.J. (1986). Cell-specific enhancers in the rat exocrine pancreas. Proc. Natl. Acad. Sci. USA *83*, 3599–3603.

Brattsand, M., and 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*, 30033–30040.

Chapdelaine, P., Paradis, G., Tremblay, R.R., and Dube, J.Y. (1988). High level of expression in the prostate of a human glandular kallikrein mRNA related to prostate-specific antigen. FEBS Lett. *236*, 205–208.

Chapdelaine, P., Gauthier, E., Ho-Kim, M.A., Bissonnette, L., Tremblay, R.R., and Dube, J.Y. (1991). Characterization and expression of the prostatic arginine esterase gene, a canine glandular kallikrein. DNA Cell. Biol. *10*, 49–59.

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.A. (1989). The glandular kallikrein family of enzymes: tissue-specific expression and hormonal regulation. Endocr. Rev. *10*, 393–419.

Clements, J.A. (1994). The human kallikrein gene family: a diversity of expression and function. Mol. Cell. Endocrinol. *99*, C1–6.

Clements, J.A. (1997). In: The kinin system, S. Farmer ed. (New York, USA: Academic Press, Inc.), 71–97.

Cleutjens, K.B., van der Korput, H.A., Ehren-van Eekelen, C.C., Sikes, R.A., Fasciana, C., Chung, L.W., and Trapman, J. (1997). A 6-kb promoter fragment mimics in transgenic mice the prostate-specific and androgen-regulated expression of the endogenous prostate-specific antigen gene in humans. Mol. Endocrinol. *11*, 1256–1265.

Diamandis, E.P., Yousef, G.M., Luo, I., Magklara, I., 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.A., Ashworth, L., Yoshida, S., Egelrud, T., Nelson, P.S., Shiosaka, S., Band, V., Little, S., Lilja, H., Stenman, U.-H., and Wain, H. (2000b). New nomenclature for the kallikrein gene family. Clin. Chem. *46*, 1855–1858.

Dihanich, M., and Speiss, M. (1994). A novel serine proteinase-like sequence from human brain. Biochim. Biophys. Acta. *1218*, 225–228.

Evans, B.A., Drinkwater, C.C., and Richards, R.I. (1987). Mouse glandular kallikrein genes. Structure and partial sequence analysis of the kallikrein gene locus. J. Biol. Chem. *262*, 8027–8034.

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., and Richards, R.I. (1988). Structure and chromosomal localization of the human renal kallikrein gene. Biochemistry *27*, 3124–3129.

Foussias, G., Yousef, G.M., and Diamandis, E.P. (2000). Identification and molecular characterization of a novel member of the siglec family (SIGLEC9). Genomics *67*, 171–178.

Frenette, G., Tremblay, R.R., Lazure, C., and Dube, J.Y. (1997). Prostatic kallikrein hK2, but not prostate-specific antigen (hK3), activates single-chain urokinase-type plasminogen activator. Int. J. Cancer *71*, 897–899.

Frydenberg, M., Stricker, P.D., and Kaye, K.W. (1997). Prostate cancer diagnosis and management. Lancet *349*, 1681–1687.

Fukushima, D., Kitamura, N., and Nakanishi, S. (1985). Nucleotide sequence of cloned cDNA for human pancreatic kallikrein. Biochemistry *24*, 8037–8043.

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

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 tumour suppressor. Cancer Res. *58*, 4782–4786.

Hansson, L., Stromqvist, M., Backman, A., Wallbrandt, P., Carlstein, A., and Egelrud, T. (1994). Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase. J. Biol. Chem. *269*, 19420–19426.

Harvey, T.J., Hooper, J.D., Myers, S.A., Stephenson, S., Ashworth, L.K., and Clements, J.A. (2000). Tissue-specific expression patterns and fine mapping of the human kallikrein *(KLK)* locus on proximal 19q13.4. J. Biol. Chem. 275, 37397–37406.

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.

Howles, P.N., Dickinson, D.P., DiCaprio, L.L., Woodworth-Gutai, M., and Gross, K.W. (1984). Use of a cDNA recombinant for the gamma-subunit of mouse nerve growth factor to localize members of this multigene family near the TAM-1 locus on chromosome 7. Nucleic Acids Res. *12*, 2791–2805.

Hu, J.C., Zhang, C., Sun, X., Yang, Y., Cao, X., Ryu, O., and Simmer, J.P. (2000). Characterization of the mouse and human PRSS17 genes, their relationship to other serine proteases, and the expression of PRSS17 in developing mouse incisors. Gene *251*, 1–8.

Kumar, A., Mikolajczyk, S.D., Goel, A.S., Millar, L.S., and Saedi, M.S. (1997). Expression of pro form of prostate-specific antigen by mammalian cells and its conversion to mature, active form by human kallikrein 2. Cancer Res. *57*, 3111–3114.

Kurtz, T.W., and St. Lezin, E.M. (1992). Gene mapping in experimental hypertension. J. Am. Soc. Nephrol. *3*, 28–34.

Lilja, H. (1985). A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. J. Clin. Invest. *76*, 1899–1903.

Little, S.P., Dixon, E.P., Norris, F., Buckley, W., Becker, G.W., Johnson, M., Dobbins, J.R., Wyrick, T., Miller, J.R., Mac Kellar, W., Hepburn, D., Corvalan, J., McClure, D., Liu, X., Stephenson, D., Clemens, J., and Johnstone, E.M. (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 transcripts: 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.

Lovgren, J., Rajakoski, K., Karp, M., Lundwall, A., and Lilja, H. (1997). Activation of the zymogen form of prostate-specific antigen by human glandular kallikrein 2. Biochem. Biophys. Res. Commun. *238*, 549–555.

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.

MacDonald, R.J., Southard-Smith, E.M., and Kroon, E. (1996). Disparate tissue-specific expression of members of the tissue kallikrein multigene family of the rat. J. Biol. Chem. *271*, 13684–13690.

Mitsui, S., Yamada, T., Okui, A., Kominami, K., Uemura, H., and Yamaguchi, N. (2000). A novel isoform of a kallikrein-like protease, TLSP/hippostasin, (PRSS20), is expressed in the human brain and prostate. Biochem. Biophys. Res. Commun. *272*, 205–211.

Murray, S.R., Chao, J., Lin, F.K., and Chao, L. (1990). Kallikrein multigene families and the regulation of their expression. J. Cardiovasc. Pharmacol. *15*, S7–16.

Murtha, P., Tindall, D.J., and Young, C.Y. (1993). Androgen induction of a human prostate-specific kallikrein, hKLK2: Characterization of an androgen response element in the 5' promoter region of the gene. Biochemistry *32*, 6459–6464.

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.

Rae, F., Bulmer, B., Nicol, D., and Clements, J. (1999). The human tissue kallikreins (KLKs 1–3) and a novel KLK1 mRNA transcript are expressed in a renal cell carcinoma cDNA library. Immunopharmacology *45*, 83–88.

Recker, F., Kwiatkowski, M.K., Piironen, T., Pettersson, K., Lummen, G., Wernli, M., Wiefelsputz, J., Graber, S.F., Goepel, M., Huber, A., and Tscholl, R. (1998). The importance of human glandular kallikrein and its correlation with different prostate specific antigen serum forms in the detection of prostate carcinoma. Cancer *83*, 2540–2547.

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., Klaassen, P., van der Korput, J.A., Geurts van Kessel, A., Romijn, J.C., and Trapman, J. (1989). The prostate-specific antigen gene and the human glandular kallikrein-1 gene are tandemly located on chromosome 19.FEBS Lett. *247*, 123–126.

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., Suurmeijer, L., Cleutjens, C.B., and Trapman, J. (1992). Characterization of the human kallikrein locus. Genomics *14*, 6–11.

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.

Southard-Smith, M., Pierce, J.C., and MacDonald, R.J. (1994). Physical mapping of the rat tissue kallikrein family in two gene clusters by analysis of P1 bacteriophage clones. Genomics *22*, 404–417.

Stephenson, S.A., Verity, K., Ashworth, L.K., and Clements, J.A. (1999). Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3–13.4. J. Biol. Chem. *274*, 23210–23214.

Sun, Z., Pan, J., and Balk, S.P. (1997). Androgen receptor-associated protein complex binds upstream of the androgen-responsive elements in the promoters of human prostate-specific antigen and kallikrein 2 genes. Nucleic Acids Res. *25*, 3318–3325.

Swift, G.H., Craik, C.S., Stary, S.J., Quinto, C., Lahaie, R.G., Rutter, W.J., and MacDonald, R.J. (1984). Structure of the two related elastase genes expressed in the rat pancreas. J. Biol. Chem. *259*, 14271–14278.

Takayama, T.K., Fujikawa, K., and Davie, E.W. (1997). Characterization of the precursor of prostate-specific antigen. Activation by trypsin and by human glandular kallikrein. J. Biol. Chem. *272*, 21582–21588.

Tanaka, T., Isono, T., Yoshiki, T., Yuasa, T., and Okada, Y. (2000). A novel form of prostate-specific antigen transcript produced by alternative splicing. Cancer Res. *60*, 56–59.

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.

Tschesche, H., Kohnert, U., Fedrowitz, J., and Oberhoff, R. (1989). Tissue kallikrein effectively activates latent degrading metalloenzymes. Adv. Exp. Med. Biol. *247A*, 545–548.

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.

Wines, D.R., Brady, J.M., Pritchett, D.B., Roberts, J.L., and MacDonald, R.J. (1989). Organization and expression of the rat kallikrein gene family. J. Biol. Chem. *264*, 7653–7662.

Wines, D.R., Brady, J.M., Southard-Smith, E.M., and MacDonald, R.J. (1991). Evolution of the rat kallikrein gene family: Gene conversion leads to functional diversity. J. Mol. Evol. *32*, 476–492.

Wolf, D.A., Schulz, P., and Fittler, F. (1992). Transcriptional regulation of prostate kallikrein-like genes by androgen. Mol. Endocrinol. *6*, 753 – 762.

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. (1998a). Sequence analysis and expression of human neuropsin cDNA and gene. Gene *213*, 9 – 16.

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

Young, C.Y., Seay, T., Hogen, K., Charlesworth, M.C., Roche, P.C., Klee, G.G., and Tindall, D.J. (1996). Prostate-specific human kallikrein (hK2) as a novel marker for prostate cancer. Prostate Suppl. *7*, 17 – 24.

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., Luo, L.Y., and Diamandis, E.P. (1999a). Identification of novel human kallikrein-like genes on chromosome 19q13.3–q13.4. Anticancer Res. *19*, 2843 – 2852.

Yousef, G.M., Luo, L.Y., Scherer, S.W., Sotiropoulou, G., and Diamandis, E.P. (1999b). 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. (1999c). 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.

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., Chang, A., Scorilas, A., and Diamandis, E.P. (2000a). Genomic organization of the human kallikrein gene family on chromosome 19q3.3 – q13.4. Biochem. Biophys. Res. Commun. *276*, 125 – 133.

Yousef, G.M., Chang, A., and Diamandis, E.P. (2000b). 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*, 11891 – 11898.

Yousef, G.M., Scorilas, A., and Diamandis, E.P. (2000c). Genomic organization, mapping, tissue expression, and hormonal regulation of trypsin-like serine protease (TLSP PRSS20), a new member of the human kallikrein gene family. Genomics *63*, 88 – 96.

Yousef, G.M., Scorilas, A., Jung, K., Ashworth, L.K., and Diamandis, E.P. (2000d). Molecular cloning of the human kallikrein 15 gene (KLK15): upregulation in prostate cancer. J. Biol. Chem., in press.

Yousef, G.M., Scorilas, A., Magklara, A., Soosaipillai, A., and Diamandis, E.P. (2000e). 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. Gene *254*, 119 – 128.

Yu, H., Diamandis, E.P., Levesque, M., Giai, M., Roagna, R., Ponzone, R., Sismondi, P., Monne, M., and Croce, C.M. (1996). Prostate specific antigen in breast cancer, benign breast disease and normal breast tissue. Breast Cancer Res. *40*, 171 – 178.