J. Biochem. 124, 491-502 (1998)

Evolution of the Trappin Multigene Family in the Suidae[1]

Yutaka Furutani,* Akira Kato,*,[2] Hiroshi Yasue,† Leeson J. Alexander,‡,[3] Craig W. Beattie,‡,[3] and Shigehisa Hirose*,[4]

*Department of Biological Sciences, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midoriku, Yokohama 226-8501; †Animal Genome Research Group, National Institute of Animal Industry, Tsukuba, Ibaraki 305-0902; and ‡U.S. Meat Animal Research Center (MARC), Clay Center, Nebraska 68933, USA

Received for publication, February 16, 1998

Trappins are a group of secretory proteins containing a WAP motif with an anchoring domain. Previous studies showed that their genes, especially those of pig, have undergone rapid evolution, which produced trappins with a broad spectrum of actions. To understand the evolution of such a rapidly evolving multigene family, we isolated trappin genes of the Artiodactyla, including pig, wart hog, collared peccary, hippopotamus, and cow, by means of polymerase chain reaction (PCR). Two genes newly isolated from wart hog are orthologs of trappin-1 (SPAI) and trappin-2 (elafin), the others are novel members of the trappin family and named trappins-6 to 11. The divergence of the sequences is greatest in the region that encodes the reactive site, and intron sequences appear to be more highly conserved than the protein-coding sequences, especially among the pig paralogs. Phylogenetic analysis showed that the trappin multigene family members of pig were generated through gene duplication after the divergence of the Suidae (pig and wart hog) and Tayassuidae (collared peccary). Similarities in the gene structure with seminal vesicle clotting proteins (REST) and WAP motif-containing proteins suggest that trappins are naturally occurring fusion proteins created through exon shuffling between ancestral REST and WAP motif-coding genes.

Key words: evolution, exon shuffling, multigene family, protease inhibitor, SINE.

Trappins are a family of proteins comprising two characteristic domains: the N-terminal transglutaminase substrate (TGS) domain and the C-terminal WAP motif region. We term the TGS domain the "cementoin moiety" since it is covalently cross-linked to extracellular matrix proteins through the action of transglutaminase and helps to anchor the biologically active WAP motif region at appropriate sites (*1-7*). The TGS domain or cementoin moiety consists of repetitive sequences rich in Gln and Lys. In this respect, it resembles the transglutaminase substrate proteins in the seminal vesicle, such as guinea pig seminal vesicle clotting proteins (SVP) (*8*) and the human semenogelins (*9, 10*). The WAP motif region has a compact structure that is formed through four disulfide bonds and is resistant to proteolysis. Members of the trappin family include SKALP/elafin (*4, 7, 11*), an elastase inhibitor, and SPAI (*6, 12*), which was initially isolated as a Na^+,K^+-ATPase inhibitor, but its true physiological function remains to be established. Recently, Zeeuwen *et al.* (*3*) coined the term, trappin, for this family of proteins and proposed a naming system according to the order of discovery (trappin-1, trappin-2, *etc.*) to allow extension of the nomenclature. The name trappin accounts for the structural and functional features of the family members; for example, it implies (i) the domain structure (consisting of the transglutaminase substrate domain and the WAP inhibitor domain), and (ii) their unique mode of action (getting trapped at the site of action), and function (trapping and inactivation of proteinases). In this paper, therefore, we use the term trappin-1 for SPAI and trappin-2 for SKALP/elafin. Trappin-1 (SPAI) is synthesized in the enteroendocrine cells in the crypt of the small intestine and then secreted into the circulation (*12, 13*). Trappin-2 is abundantly expressed in

[1] This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan, and by research grants from the Proposal-based Advanced Industrial Technology R&D Program of NEDO, SRF, the Cosmetology Research Foundation, and the Life Insurance Association of Japan. The nucleotide sequence(s) reported in this paper has been submitted to the DDBJ/GenBank™/EMBL Data Bank under accession number(s) AB003281-AB003285, D50322, D50323, D83668, and AB011010.

[2] Supported by a Research Fellowship for Young Scientists from the Japan Society for the Promotion of Science. Present address: Molecular Cell Science Laboratory, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074.

[3] Present address: Veterinary PathoBiology, University of Minnesota, St. Paul, Minnesota 55109, USA.

[4] To whom correspondence should be addressed. Tel: +81-45-924-5726, Fax: +81-45-924-5824, E-mail: shirose@bio.titech.ac.jp

Abbreviations: TGS, transglutaminase substrate; WAP, a protein motif first characterized in whey acidic protein; trappin, transglutaminase substrate domain and WAP motif-containing protein; REST, rapidly evolving seminal vesicle transcribed; SVP, seminal vesicle secretory protein(s); SKALP, skin-derived antileukoproteinase; SPAI, WAP motif protein with weak sodium/potassium ATPase inhibitory activity; SLPI, secretory leukocyte protease inhibitor or antileukoproteinase; PCR, polymerase chain reaction; bp, base pair(s); kb, kilobase(s) or kilobase pair(s); SINE, short interspersed repetitive element.

the trachea and skin (*2, 3, 14, 15*). The cDNAs and genomic clones of the members of the trappin family have been isolated from various mammalian species such as: trappin-2 from man (*4, 16*), monkey (*3*), pig (*14*), and cow (*3*); trappins-1 and -3 from pig (*6, 14*); and trappin-4 and -5 from cow (*3*). The numbers of trappin genes vary depending on the species of the animals (*3, 14*).

Trappins are unique not only in their mode of action, but also in their gene structure. Trappin genes are composed of three exons and two introns. Exon 1 encodes the signal peptide; exon 2, the mature protein sequence; and exon 3, the 3′-noncoding region (*14, 16*). Our previous analysis of pig trappin genes indicated that they are very similar in size and exon-intron organization, and form a multigene family that contain a SINE termed PRE-1 in intron 2 (*14*). Sequence divergence is greatest in mature protein-coding exon 2 rather than noncoding exons and introns. The latter are remarkably conserved (*14*). The rapidly evolved exon 2 and the presence of a SINE, which is not present in the trappin genes of the other animals, are interesting from an evolutionary point of view. We therefore extended analysis of the trappin gene family to the Artiodactyla to obtain clues as to the rapid evolution of this multigene family in the Suidae. We also formulated the hypothesis that exon 2 of the trappin genes evolved through exon shuffling and fusion between the genes for transglutaminase substrate proteins such as seminal vesicle clotting proteins (*8, 17*), and WAP motif proteins such as secretory leukocyte protease inhibitor (SLPI) (*18*).

EXPERIMENTAL PROCEDURES

DNA Samples—Genomic DNAs were isolated from pig (*Sus scorfa*), wart hog (*Phacochoerus africanus*), collared peccary (*Pecari tajacu*), hippopotamus (*Hippopotamus amphibius*), and bovine (*Bos taurus*) tissues as described (*19*). The first strand cDNA used for the isolation of cDNA for pig trappin-7 was prepared from mRNA of pig small intestine as previously described (*14*).

Isolation of Trappin Family Genes—Seven PCR primers were designed based on the results of sequence comparison among human trappin-2 (elafin), pig trappin-1 (SPAI), and pig trappin-2 (elafin). The sequences of the primers used for cloning were as follows: primer PS1, 5′-CAAGGGAACCCTGATCCACTAGAC-3′; primer PS2, 5′-GTCTAAATGTGGGGATGTTTCTTC-3′; primer PS3, 5′-TCCCCAGTGAGGTGAGAACT-3′; primer PA1, 5′-GAGGACCAGAGATGTGATTCAGCA-3′; primer PA2, 5′-ACCAAGTCCATGTGTCATCCCTCT-3′; primer PS980-1, 5′-GAGAATTCACCTTCCTGACAACATGAG-3′; and primer PA980-3, 5′-TTCTGCAGGTGCAGCAAGACTCCAC-3′. The primer sites are shown in Fig. 1A. PCR was carried out with genomic DNA digested with *Eco*RI or first strand cDNA as template and an Expand Long Template PCR System (Boehringer Mannheim, Germany). Primers PS1 and PA1 were used to amplify the pig, wart hog, and collared peccary trappin genes. Primer sets (PS2 and PA2; PS3 and PA1) were used to amplify fragments of the hippopotamus trappin gene. PS980-1 and PA980-3 were used to amplify a fragment of cDNA for pig trappin-7. PCR reactions were performed for 30 cycles under the following conditions: denaturation for 30 s at 95°C, annealing for 1 min at 45°C, and extension for 2 min at 68°C. The final PCR products were fractionated by 1.2% agarose gel electrophoresis and then transferred to nitrocellulose filters (Schleicher & Schüll, Dassel, Germany). The filters were prehybridized for 1 h at 42°C in 20% formamide, 6×SSPE (1×SSPE is 0.15 M NaCl, 15 mM NaH_2PO_4, pH 7.0, 1 mM EDTA), 1% SDS, and 5×Denhardt's solution (1×Denhardt's solution is 0.02% each of bovine serum albumin, polyvinylpyrrolidone, and Ficoll). Hybridization was performed for 12–16 h at 42°C in the same solution containing 1.0×10^6 cpm/ml of a probe labeled with a Ready-To-Go DNA labeling kit (Pharmacia, Piscataway, NJ, USA). Following hybridization, each filter was washed twice with 2×SSC (1×SSC is 0.15 M NaCl, 15 mM sodium citrate, pH 7.0) and 0.1% SDS at room temperature for 30 min, and then 0.5×SSC and 0.1% SDS at 55°C for 8 h. The 617-bp fragment of the pig trappin-1 (SPAI) gene, generated by digestion of the pig trappin-1 (SPAI) gene with *Pst*I and *Sac*I, was used as a probe for hybridization. Hybridized bands were extracted from the gel and subcloned into pBluescript II (Stratagene, La Jolla, CA, USA). The resulting clones were confirmed by colony hybridization and DNA sequencing.

DNA Sequencing—DNA sequences were determined by the dideoxynucleotide chain-termination method using a SequiTherm Long-Read Cycle Sequencing Kit-LC (Epicentre Technologies, Madison, WI, USA), and automated sequencing performed with an LI-COR LC4000 (LI-COR, Lincoln, NE, USA). At least three clones of each trappin gene were sequenced on both strands. Sequence data were organized and analyzed using the program, GENETYX-MAC (Software Development, Tokyo).

Construction of Phylogenetic Trees—The nucleotide or amino-acid sequences of trappins were aligned using the program, CLUSTAL W (*20*). Phylogentic analysis was performed using PHYLIP (Phylogenetic Inference Package, version 3.5) (*21*).

Southern Blot Analysis—Ten micrograms of DNAs (pig and collared peccary) were digested with restriction enzymes, fractionated on a 0.7% agarose gel, and then transferred to a nylon membrane. The membrane was prehybridized at 37°C in 20% formamide, 6×SSPE, 5×Denhardt's solution, 100 μg/ml denatured herring sperm DNA, and 0.5% SDS for 3 h, and then hybridized with ^{32}P-labeled hippopotamus trappin gene fragment (727-bp PCR product) at 37°C in 20% formamide, 6×SSPE, 5×Denhardt's solution, 100 μg/ml denatured herring sperm DNA, and 0.5% SDS for 16 h. The membrane was washed two times at 55°C with 1×SSC and 0.1% SDS for 1 h, exposed to an imaging plate for 12 h, and then analyzed with a Bioimage Analyzer (model BAS 2000; Fuji Film, Tokyo).

Isolation of Yeast Artificial Chromosomes (YACs) Containing the Pig Elafin (Trappin-2) and SLPI Genes—Clones containing the pig elafin gene were identified by PCR-based screening (*22*) of a pig YAC library (*23*). The PCR assay utilized the primers, 5′-GGGATGTTTCTTCTTTCCACAGGT-3′ and 5′-TTGGTGAGTAAGATTAGACGTTTG-3′, which correspond to the sequences of intron 1 and exon 2 of the pig elafin gene, respectively. The isolation of SLPI genes within these YACs was performed by PCR using the primers, 5′-TGCCCTCCTAGAAAAATTGTCCAGTG-3′ and 5′-GCGCAAGTATCTCGGCAACATTTC-3′, which correspond to sequences of exon 2 of the pig SLPI gene. The existence of other trappin genes within these

YACs were also examined by PCR using the primers, 5′-CAAGGGAACCCTGATCCACTAGAC-3′ and 5′-ACCAAGTCCATGTGTCATCCCTCT-3′, which correspond to highly conserved regions in intron 1 and intron 2 of pig trappin genes.

RESULTS

Cloning and Sequence Analysis of Trappin Genes of the Artiodactyla—Using the primers shown in Fig. 1A, we cloned trappin-coding regions (exon 2) and nearby intron sequences by PCR. To avoid possible PCR artifacts, three to six clones of each gene were sequenced and the consensus sequences were used for the analysis. The error frequency due to nucleotide misincorporation was very low (0.01–0.06%). No chymeric or recombinant products were found. Two genes for wart hog trappins and one gene each for peccary, hippopotamus and bovine trappins were isolated and characterized (Fig. 1B). Three additional members of the pig trappin gene family were also identified by analyzing the pig genome and cDNA. The nucleotide sequences and deduced amino acid sequences are shown in Fig. 2. Comparison of the amino acid sequences of the newly isolated genes with those of previously determined trappins [*i.e.* trappin-1, -2, -3, -4, and -5 (*3, 14*)] revealed that two trappin genes from wart hog are orthologs of trappin-1 and -2, and the others are new members. The novel trappins have been named here trappin-6, -7, -8, -9, -10, and -11 according to the nomenclature suggested by Zeeuwen *et al.* (*3*).

Previous analysis of genes for pig trappins (trappin-1, -2 and -3) by Tamechika *et al.* (*14*) showed the presence of a SINE (PRE-1) in intron 2, which is expected to have played a role in the evolution of the trappin multigene family in pig. In the case of the newly isolated genes, the wart hog (trappin-1 and -2), pig (trappin-7, -8, and -9), and collared peccary (trappin-10) genes contain a SINE (Figs. 1B and 2) like the three pig trappin genes previously isolated (*14*); the hippopotamus gene (trappin-11) lacks a SINE in intron 2 (Fig. 1B), as does the human trappin gene (*14*), suggesting that the insertion of a SINE into the trappin gene occurred after the divergence of the Suina (Fig. 1C).

Phylogenetic Analysis of Trappin Genes of the Artiodactyla and the Paralogous Origin of the Trappin Multigenes in the Suidae—To clarify the phylogenetic relationship of these trappin genes, we constructed a phylogenetic tree by the neighbor-joining method (*24*) (Fig. 3). The tree shows that the Suidae (wart hog and pig) sequences are clustered on the same branch of the tree. The Suidae branch consists of two subgroups: The first group contains trappin-1, -3, -7, and -9, and the second group contains trappin-8, and the pig and wart hog orthologs of trappin-2. The position of trappin-10 of collared peccary suggests that the Suidae trappin paralogs originated from their ancestral gene through duplications of the genes after the divergence of the Tayassuidae and Suidae [16.8 million years ago (*25*)]. The occurrence of wart hog orthologs for pig trappin-1 and -2 also suggested that the duplications occurred before the divergence of wart hog and pig [1.4 million years ago (*25*)]. The phylogeny of the subgroups of these Suidae (trappin-1, -3, -7, -8, and -9, and the pig and wart hog orthologs of trappin-2) paralogs shows the second and third duplications occurred during the evolution of the trappin multigenes in the Suidae. The bovine sequences determined by Zeeuwen *et al.* (*3*) (trappin-4 and -5) and us (trappin-6) are also clustered on the same branch, suggesting these bovine paralogs (trappin-4, -5, and -6) occurred after bovine speciation.

Genomic Southern Blot Analysis—Transposable elements are generally believed to be a strong force generating variations of genes (for a recent example, see Ref. *26*), and pig trappin genes contain a SINE (PRE-1) in intron 2 (*14*). Sequence data revealed that the insertion of PRE-1 only

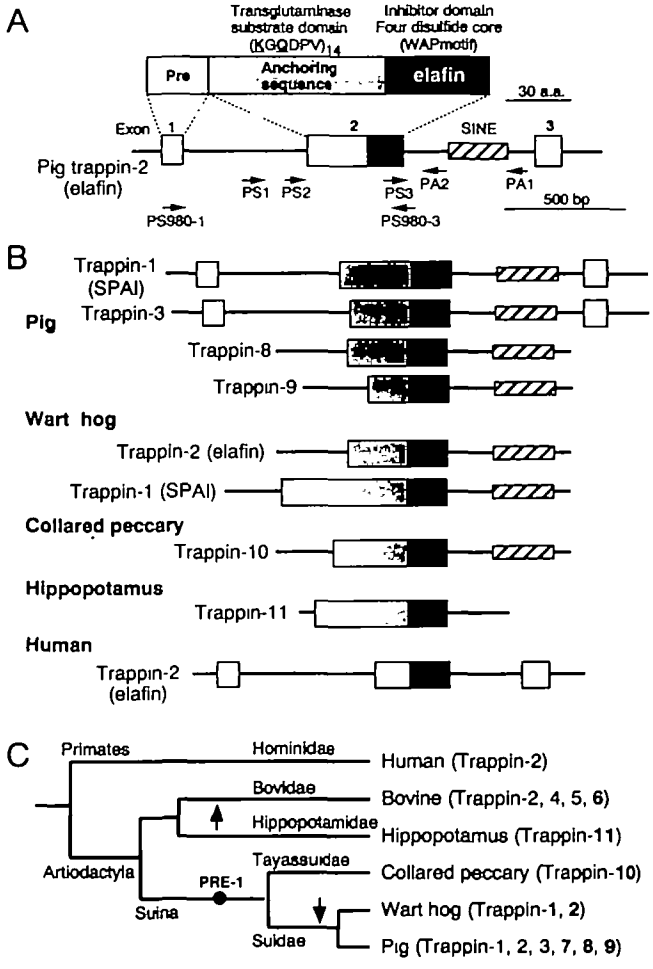


Fig. 1. **Trappin gene family.** A: Structural features of pig trappin-2 (SKALP/elafin) and its genomic organization. Exons are represented by boxes and introns by lines. The SINE is indicated as a small striped box. The nucleotide sequence of exon 1 encodes a signal peptide. The nucleotide sequence of exon 2 encodes a transglutaminase substrate domain (gray box) and an inhibitor domain (solid box). Exon 3 is the 3′-noncoding region. Arrows indicate primers. B: Complete or partial structures of trappin genes. C: Phylogeny of the Artiodactyla. A phylogenetic tree was constructed according to the results of a recent study on the phylogeny of the Artiodactyla by Shimamura *et al.* (*47*). The names of the orders (Primates and Artiodactyla) and families (Hominidae, Bovidae, Hippopotamidae, Tayassuidae, and Suidae) are shown with the tree. Suina is the name of the infraorder that consists of the Suidae and Tayassuidae. The black circle indicates the insertion of PRE-1 into the trappin genes in the Suina. Possible times of duplication of the trappin genes are indicated by arrows. Trappins that have been isolated and characterized from each animal are shown. The new members of the family which were isolated in this study are indicated in bold letters.

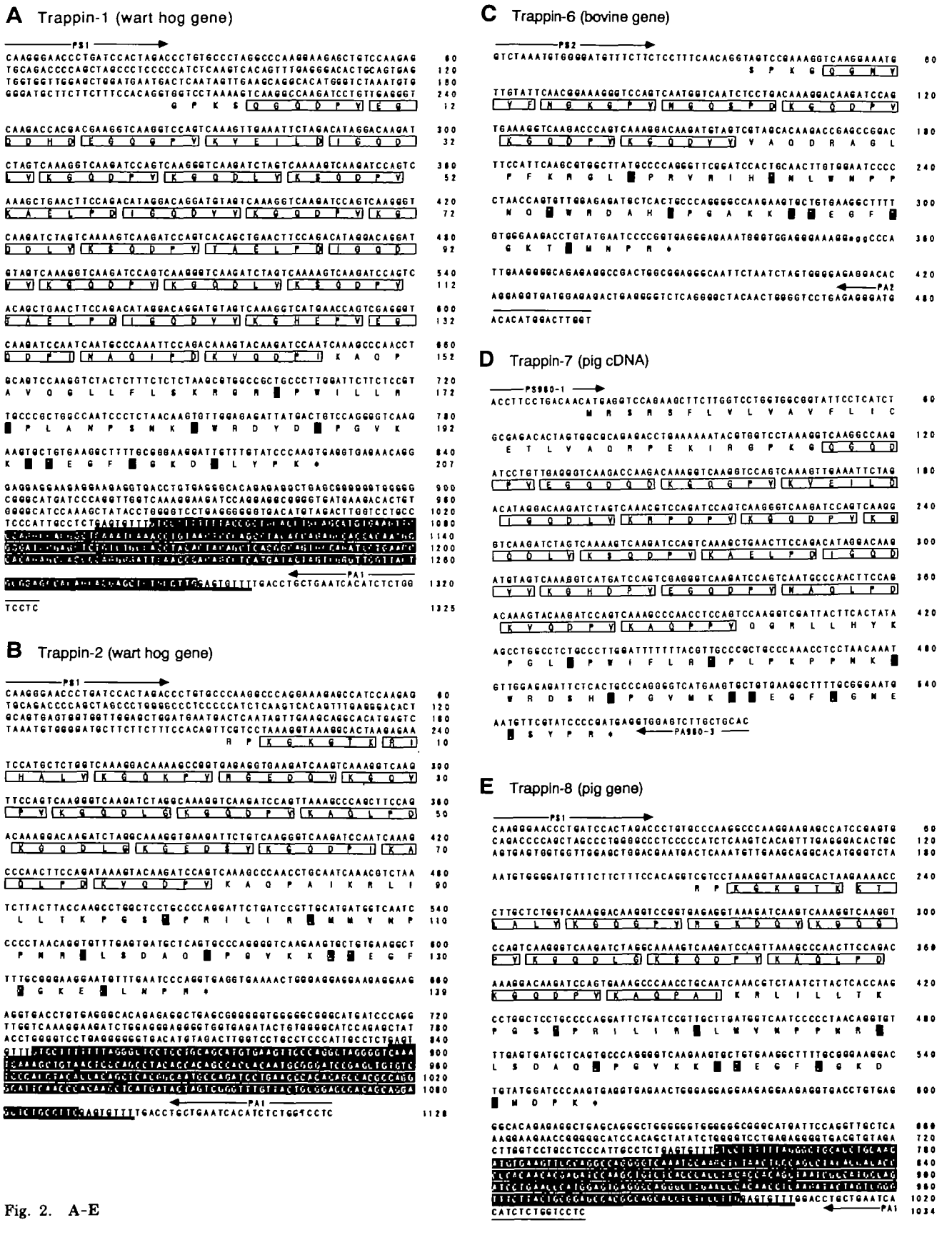
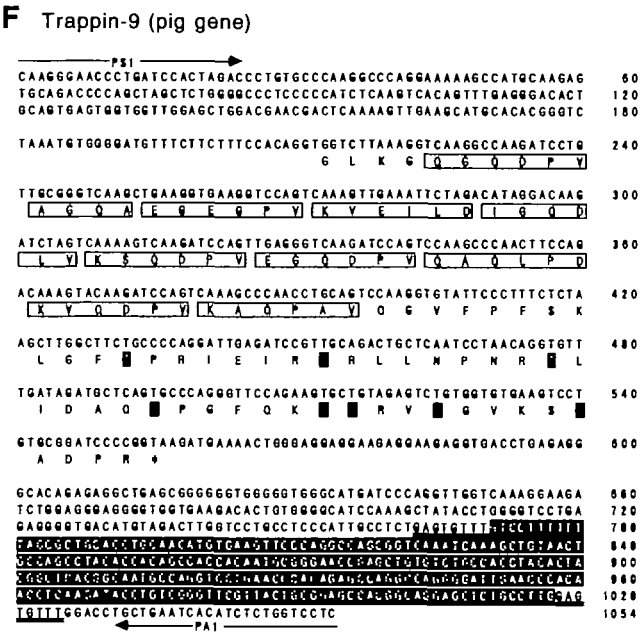
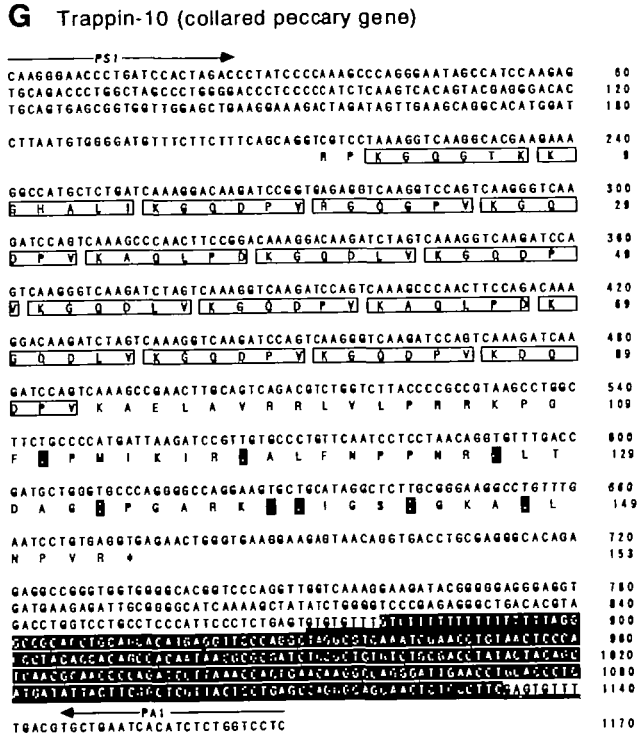


Fig. 2. A-E

occurred in the trappins of the Suidae and Tayassuidae. However, the phylogenetic analysis showed the duplication of the pig trappin genes occurred after the divergence of the Suidae and Tayassuidae. Furthermore, only a single trappin gene (trappin-10) was isolated by PCR from genomic DNA of collared peccary. To determine if multiple trappin genes are also present in peccary, genomic DNA of collared peccary was analyzed by Southern blotting. Compared to the pig genome, a simple pattern was observed for collared peccary (Fig. 4, left), consistent with the assumption of a single gene or a small number of genes for trappin. These data suggested that PRE-1 alone was not sufficient for the duplication and rapid evolution of trappin paralogs in the Suidae.

Conservation of Intron Sequences among the Paralogous Trappin Genes in the Suidae—Figure 5 shows the homology of intron 1, exon 2, and intron 2 of trappins of the Artiodactyla. As previously noted (*14*), the similarities in intron sequences are remarkably high compared to those in the exon 2 sequences among the trappin paralogs. This





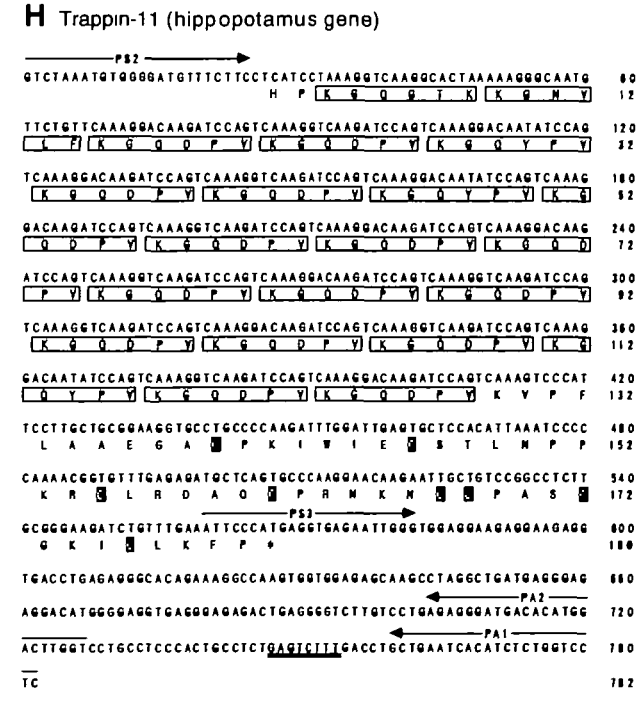


Fig. 2. **Nucleotide sequences of genes and cDNAs of trappins, and the deduced amino acid sequences.** The primers used for PCR amplifications are indicated by arrows above the nucleotide sequence. Conserved Cys residues within WAP motifs are indicated by white on black. Lys- and Gln-rich repeated amino acid sequences are boxed. SINEs (PRE-1) are boxed and shaded, which are flanked by a pair of direct repeats marked by double underlines. A homologous sequence of the direct repeat is also found in the hippopotamus trappin gene that has no SINE, and is indicated by a double underline.

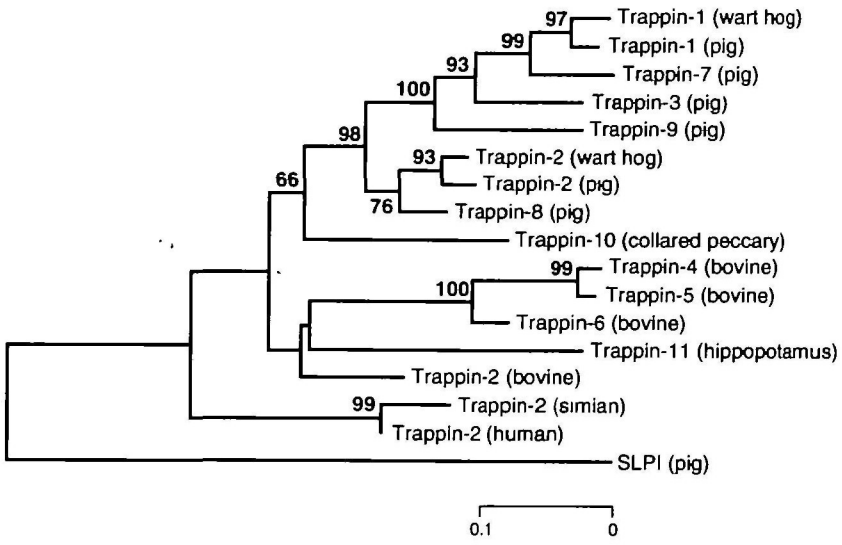


Fig. 3. **Phylogeny of trappin genes.** A phylogenetic tree was constructed by comparing the nucleotide sequences of trappin gene exon 2 (280 bp) using the neighbor-joining method. The sequence of the second WAP motif of pig SLPI was used as an outgroup. The numbers for interior branches refer to the bootstrap values for 100 replications. Bootstrap values of less than 50% are not given. The scale at the bottom is units of nucleotide substitutions per site.

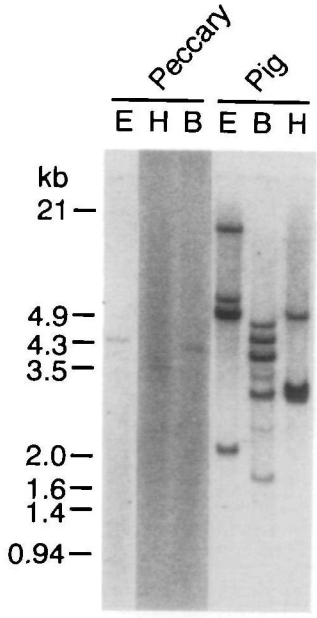


Fig. 4. **Southern hybridization analysis of genomic DNAs of pig and collared peccary.** Five micrograms of genomic DNAs was digested with restriction enzymes (*Eco*RI, *Bam*HI, and *Hin*dIII), fractionated on a 0.7% agarose gel, and then blotted onto a nylon membrane. The membrane was probed with a radiolabeled fragment of the hippopotamus trappin gene. The migration of size markers is shown on the left of the figure. The generation of a multigene family was only detected in the Suidae (pig), *i.e.* not in the Tayassuidae (collared peccary).

tendency was remarkable among the paralogs of the Suidae, suggesting the homogenization of intron sequences occurred during the evolution of the trappin multigenes in the Suidae, possibly through gene conversion.

Amino Acid Sequence Comparison in the WAP Motif or the Inhibitor Domain—Figure 6 shows the deduced amino acid sequences of the WAP motif regions of Artiodactyla trappins as well as those previously determined for the trappins from pig (*6, 14*), cow (*3*), man (*4, 16*), and monkey (*3*). The eight cysteine residues (indicated by *) that form the four disulfide core of the WAP motif are all conserved among the trappin family members. Two regions, V1 and V2, are quite variable, one of which (V2) has been demonstrated, by X-ray crystallography, to be the inhibitory site that interacts with the active site of a proteinase (*27, 28*). As typically seen in pig, the generation, through gene duplication and divergence, of multiple trappin family members with different inhibitor selectivities appears to be a very useful mechanism for protecting tissues and cells from attack by harmful proteinases.

Diversification of the Cementoin-Like Transglutaminase Substrate Domain—The transglutaminase substrate domain (TGS) of trappins consists of semiconserved repetitive sequences of 6 amino acid residues (*2*). The number of repeats was found to vary depending on the species and the family member; for example, wart hog trappin-1 has a relatively long TGS domain of 25 repeats (Fig. 7) and human trappin-2 has a short TGS domain of only 5 repeats (*2*). Alignment of the nucleotide sequences corresponding to the TGS domains of trappins of the Artiodactyla (Fig. 7) allowed the construction of an evolutionary tree that provides a reasonable explanation of how the present day complex structures of the trappin TGS domains evolved (Fig. 8).

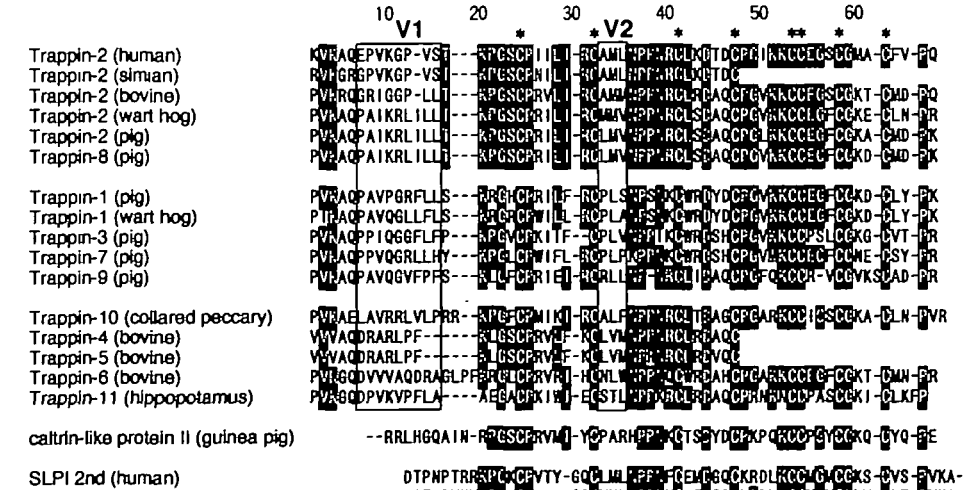
Strong Similarity between Trappins and the Second WAP Domain of Secretory Leukocyte Protease Inhibitor (SLPI) Revealed by Analysis of the Phylogenetic Relationship among Trappins and Other Proteins Containing WAP Motif(s)—To determine the phylogenetic relationship of trappins with other proteins, a phylogenetic tree was constructed, using the neighbor-joining method (*24*), for the amino acid sequences of WAP domains. This analysis was also expected to help us to determine the origin of the trappin gene. The tree in Fig. 9A shows that (i) all trappins belong to a single branch and (ii) the second WAP region of SLPI is the closest to trappins. SLPI is a serine protease inhibitor consisting of two WAP domains, both of which are encoded by individual exons possibly having arisen through duplication of an WAP-coding exon. The phylogeny among the two WAP domains of SLPI and trappins suggested that the exon duplication of the SLPI gene occurred before the divergence of SLPI and trappin.

order	family	species & trappin gene	pig, 1	pig, 2	pig, 3	pig, 7	pig, 8	pig, 9	war., 1	war., 2	pec , 10	hip., 11	bov., 2	bov , 4	bov., 5	bov., 6	sim., 2
Artiodactyla	Suidae	pig, 2	***99*** (82) 88														
		pig, 3	***97*** (84) ***97***	***97*** (80) ***90***													
		pig, 7	(***90***)	(***79***)	(84)												
		pig, 8	***95*** (80) 87	***96*** (***93***) ***96***	***95*** (80) 87	(78)											
		pig, 9	***97*** (79) ***96***	***96*** (81) 88	***94*** (81) ***96***	(80)	***94*** (79) 86										
		wart hog, 1	***91*** (***95***) ***93***	***92*** (83) 87	***92*** (82) ***93***	(88)	***91*** (81) 85	89 (78) ***95***									
		wart hog, 2	***98*** (83) ***94***	***98*** (***96***) 87	***96*** (82) ***94***	(81)	***95*** (***91***) 87	***95*** (80) ***95***	***93*** (84) ***96***								
	Tayassuidae	collared peccary, 10	89 (72) 80	89 (78) 78	87 (71) 82	(73)	87 (80) 76	87 (75) 82	83 (72) 83	88 (77) 82							
	Hippopo-tamidae	hippopotamus, 11	(68) 70	(72) 75	(72) 73	(69)	(74) 69	(69) 70	(66) 68	(73) 69	(74) 71						
	Bovidae	bovine, 2	(73)	(79)	(72)	(73)	(81)	(73)	(72)	(78)	(77)	(77)					
		bovine, 4	(64)	(69)	(65)	(65)	(71)	(65)	(62)	(69)	(66)	(71)	(77)				
		bovine, 5	(65)	(69)	(64)	(64)	(71)	(65)	(62)	(68)	(66)	(71)	(75)	(***97***)			
		bovine, 6	(71) 61	(72) 55	(71) 61	(70)	(74) 55	(69) 61	(69) 60	(73) 61	(72) 62	(71) 75	(78)	(89)	(88)		
Primates	Cercopi-thecidae	simian, 2	(65)	(67)	(67)	(67)	(70)	(66)	(65)	(68)	(68)	(73)	(76)	(85)	(64)	(65)	
	Hominidae	human, 2	67 (70) 63	68 (73) 58	67 (69) 61	(71)	68 (75) 54	67 (68) 58	69 (70) 58	67 (73) 60	68 (72) 61	(73) 64	(79)	(87)	(67)	(70) 57	(***97***)

intron 1
(exon 2)
intron 2

Fig. 5. **Comparison of the sequences of intron 1, exon 2, and intron 2 for trappin genes.** Numbers indicate the identity (%) of the sequences. The identity of exon 2 is given in parenthesis. The identity of introns 1 and 2 is given over and under the identity of exon 2, respectively. Values of more than 90% are shown in bold italic letters. Abbreviations: wart hog, 1 or war., 1, wart hog trappin-1.

Fig. 6. **Comparison of the amino acid sequences of the WAP motifs of trappins.** For alignment, the amino acid sequences of the trappins and SLPI WAP motifs are shown in the single-letter code. Conserved residues among all trappins are shown in black boxes. Residues conserved among the orthologs of trappin-2 (elafin) are indicated by gray boxes. The eight Cys residues of WAP motifs are indicated by *. Highly variable regions are boxed, and numbered V1 and V2 (*14*). The accession numbers of the compared proteins are as follows: human trappin-2 (elafin), D13156; simian trappin-2, AJ-223215; pig trappin-1 (SPAI), D50302; pig trappin-2, D50319; pig trappin-3, D50321; pig trappin-7, D50323; pig trappin-8, D50322; pig trappin-9, AB003285; collared peccary trappin-10, AB003283; hippopotamus trappin-11, AB003284; bovine trappin-2, AJ223216; bovine trappin-4, AJ223217; bovine trappin-5, AJ233218; bovine trappin-6, AB011010; guinea pig caltrin-like protein II, P22075; human SLPI, X04470 and X04502; pig SLPI, M57446; and mouse SLPI, U73004.



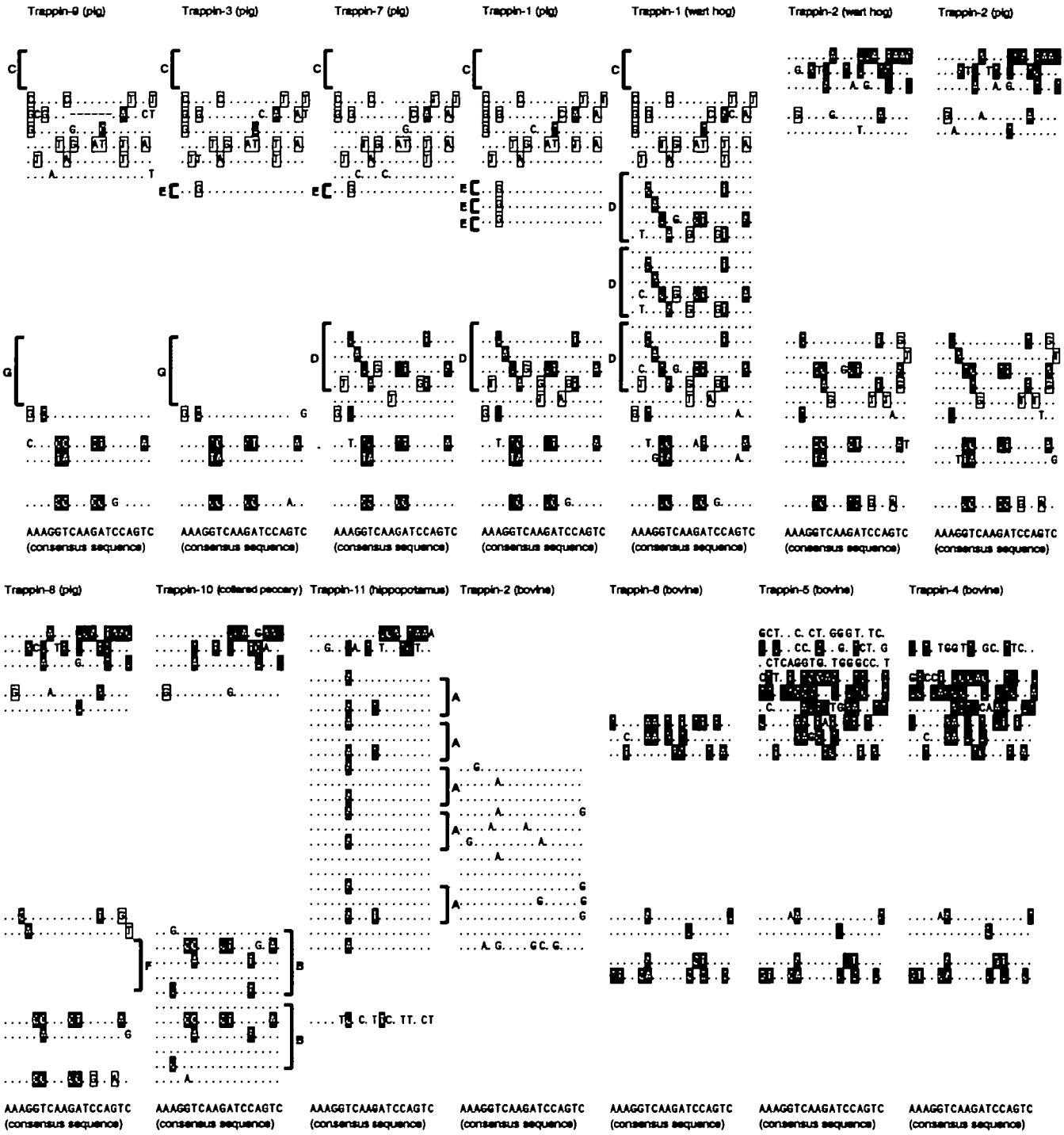


Fig. 7. **Sequence alignment of the TGS domains of trappin genes in the Artiodactyla.** Nucleotides identical to the consensus sequence are indicated by dots, and gaps are indicated by bars. The nucleotides boxed in black represent those different from the consensus sequence but conserved among the trappin genes in the Artiodactyla, which helped to define the evolution of the genes. Insertions and deletions of the repetitive sequences are indicated by blocks A-G.

A WAP-motif containing protein termed "caltrin-like protein II," which was purified from guinea pig seminal vesicles by Coronel *et al.* (*29*), belongs to the same branch of the WAP family as trappin. This fact raises the possibility that although the purified caltrin-like protein II lacked a TGS domain, it may have an amino-terminal TGS domain in its native form and therefore is a member of the trappin gene family. Similar situations, namely the loss of the TGS domain, have been encountered during the purification of trappin-1 (SPAI) and trappin-2 (SKALP/elafin) (*11, 30*).

Proximal Localization of the Trappin-2 and SLPI Genes—Many related genes that share a common ancestor are known to occur in close proximity on the same chromosome. In humans, the trappin-2 (elafin) (*31*) and SLPI

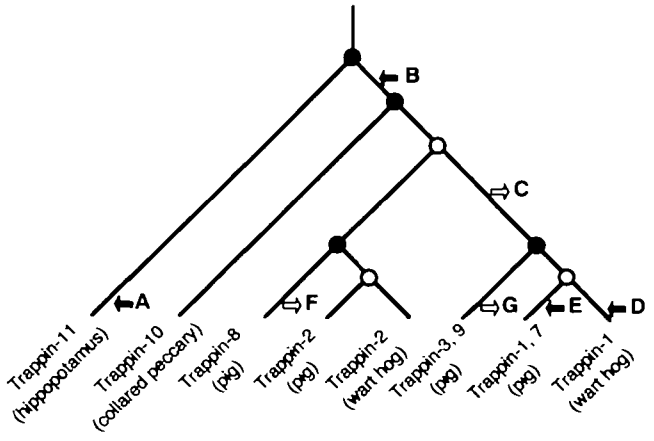


Fig. 8. **Rapid evolution of the TGS domains of trappin genes in the Artiodactyla.** Phylogeny of the TGS domains of trappin genes in the Artiodactyla is shown with the possible times of insertions and deletions of the repetitive sequences (blocks A–G, Fig. 7). Divergence of the species is shown by filled circles, and the duplication of trappin genes is shown by open circles.

[human STS WI-6969 (*32*)] genes are located on chromosome 20q12-q13 (Fig. 9C). We therefore determined whether or not this is the case for the pig genes using a YAC library. Two YAC clones that contain the pig elafin gene were first isolated by PCR screening and then examined for co-localization of the SLPI gene. PCR amplification with primers specific for the SLPI gene indicated that the YAC clones were also positive for the pig SLPI gene, demonstrating the close localization of the two genes on the pig chromosome.

DISCUSSION

In the present study, we characterized the trappin gene family in the Artiodactyla in terms of the number of family members, the nucleotide and amino acid sequence similarities, and the presence of a SINE in intron 2. This allowed us to determine the time of duplication of the trappin genes and to construct a phylogenetic tree. The amino acid sequences of Artiodactyla trappins, especially those of variable region V2, will also provide valuable information for the design of therapeutic agents such as anti-inflammatory and anti-infection drugs.

Origin of the Trappin Gene Family—Recently, Lundwall and Ulvsbäck (*33*) pointed out the structural similarity between the trappin-2 (SKALP/elafin) and REST genes, and suggested that the proteinase inhibitor, trappin-2 (SKALP/elafin), is a member of the REST gene family. Members of the REST gene family are composed of three exons and are transcribed almost exclusively in the seminal vesicles. Like in the trappin genes, exon 1 codes for the signal sequence; exon 2, the entire transglutaminase substrate protein; and exon 3, the 3′ untranslated region. Exons 1 and 3 are highly conserved among the members while the exon 2 sequences are quite variable, suggesting rapid evolution of transglutaminase substrate proteins with new specificities; thus the name rapidly evolving seminal vesicle transcribed (REST) genes (*34*). As Lundwall and Ulvsbäck (*33*) pointed out, the similarity in exons 1 and 3 extends to the members of the trappin gene family (Fig. 9). The products of the REST gene family members are called semenogelin in man (*35*), seminal vesicle secretory proteins in guinea pig [SVP (*8*)] and rat [SVS (*36*)], and semenoclotin in mouse (*17*).

Analysis by Hagstrom *et al.* (*8*) of the guinea pig REST gene (termed GP1G) revealed another interesting relationship between the trappin and REST genes. Namely, the GP1G gene encoding guinea pig SVP was shown to consist of tandem repeats of the REST genes with a total of 6 exons (exons a1, a2, a3, b1, b2, and b3; Fig. 9B), three of which (exons a2, a3, and b1) are pseudo-exons. Guinea pig SVP is therefore encoded by the combination of exons a1, b2, and b3. Furthermore, sequence comparison indicated that the first gene (a1-a2-a3), although being a pseudo-gene, is similar to the human semenogelin genes, and the second gene (b1-b2-b3) is more closely related to the human trappin-2 (SKALP/elafin) gene rather than to any of the SVP genes.

We previously determined the gene structures of trappin family members and showed that trappins are fusion proteins comprised of SVP-like transglutaminase substrate domains (termed cementoin moieties) and WAP motif proteins (*1, 2, 6*). In the present study, we showed that the WAP motif-coding regions of the trappin genes are much more similar to the second WAP motif region of the SLPI gene than those of the other genes that have WAP motif-coding region(s) (Fig. 9A). These considerations led us to propose the following mechanism for the evolution of the trappin genes (Fig. 9B): The trappin genes have evolved through duplication of the ancestral REST gene and successive shuffling of the exons encoding the WAP motif. In support of this hypothesis, the human REST genes, trappin genes, and WAP motif genes are clustered on chromosome 20, bands q12-13.1 (Fig. 9B) (*31, 32, 35*). Also in the case of pig, we mapped the genes for trappin-2 (elafin) and SLPI on the same YAC clone.

It is also likely that, among all members, trappin-2 is the prototype of the trappin gene family. This is plausible for the following two reasons: (i) In addition to the structural similarity, the functions of SLPI and trappin-2 are very similar in the inhibition of neutrophil elastase (*7, 37, 38*), and (ii) orthologs of trappin-2 are present in a variety of species such as man, monkey, cow, pig, and wart hog. Although trappin-2 genes have not yet been isolated from collared peccary and hippopotamus on a number of PCR analyses of their genomes, it is likely that these animals may also have trappin-2 genes.

Mechanism and Biological Significance of Multiplication of the Trappin Gene in the Suidae—The numbers of members of the trappin gene family vary widely depending on the species; for example, multiple trappin genes have been demonstrated in the pig whereas only a single trappin gene has been identified in the human genome (*14, 16*). Based on the fact that there is a SINE in intron 2 of the pig trappin genes but no such sequence is found in the human counterpart, we suspected that insertion of a SINE might have stimulated the duplication of the trappin genes. To examine this possibility, we analyzed Artiodactyla trappin genes, and determined the number of genes and the presence of SINE. There appears to be a relationship between the number of family members and the presence of SINE (Fig. 1), but in the case of the peccary only one trappin gene was detected despite the presence of a SINE,

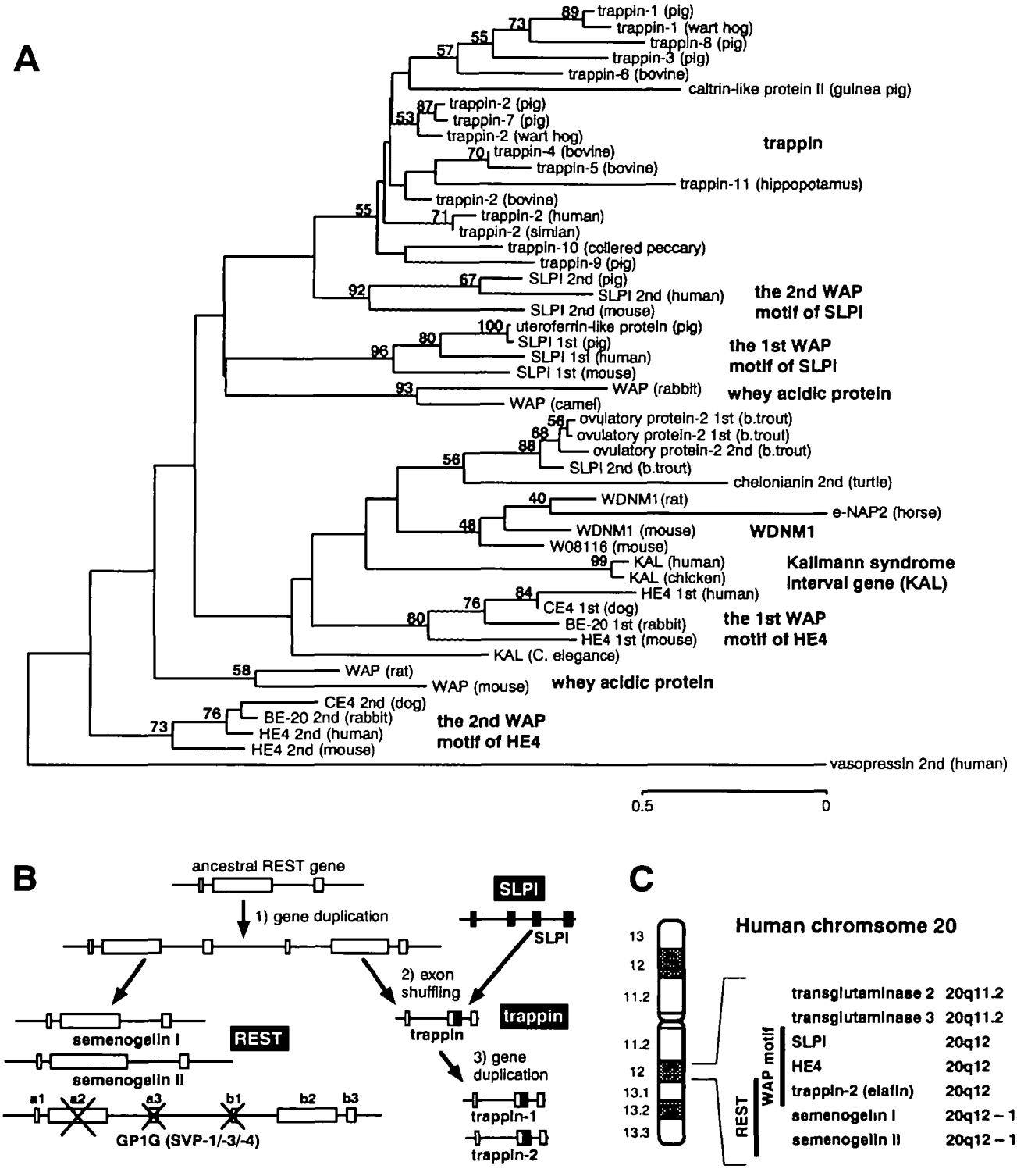


Fig. 9

suggesting that a SINE alone is not sufficient to trigger gene duplication. The repetitive sequences in exon 2 may also be a potential candidate for the *cis* element responsible for the rapid evolution of the trappin genes (*39-41*). In this context, it is interesting to note that the number of repeats is smallest in the human and peccary trappin genes, that have not been duplicated. Gene conversion, mismatch repair, and replication slippage are likely to be involved in the mechanism of this evolution.

The phylogenetic analysis shown in Fig. 3 suggests the very recent duplication of the paralogs of pig trappins after the divergence of the Suidae and Tayassuidae. The multiplication of trappin genes in the Suidae led to a highly variable protein-coding region (exon 2) (Fig. 6 and Ref. *14*)

Fig. 9. **Hypothetical mechanism of evolution of the trappin family and chromosomal locations of the related genes.** A: Phylogeny of trappins and other related proteins which have one or two WAP motifs. A phylogenetic tree was constructed by comparing the peptide sequences of WAP motifs using the neighbor-joining method. The accession numbers of the compared proteins are as follows: human trappin-2 (elafin), D13156; pig trappin-1 (SPAI), D50302; pig trappin-2, D50319; pig trappin-3, D50321; pig trappin-7, D50323; pig trappin-8, D50322; pig trappin-9, AB003285; collared peccary trappin-10, AB003283; hippopotamus trappin-11, AB003284; bovine trappin-2, AJ223216; bovine trappin-4, AJ223217; bovine trappin-5, AJ233218; bovine trappin-6, AB011010; guinea pig caltrin-like protein II, P22075; pig uterine secretory protein, M31216; human SLPI, X04470 and X04502; pig SLPI, M57446; mouse SLPI, U73004; mouse WAP (whey acidic protein), J00649; rat WAP, J00801; camel WAP, P09837; rabbit WAP, X07943; red sea turtle chelonianin, P00993; brook trout SLPI, U03890; brook trout ovulatory protein-2, U67854; mouse WDNM1, X93037; rat WDNM1, X13309; a mouse EST clone homologous to WDNM1, W08116; horse antibiotic peptide eNAP-2; human KAL (Kallmann syndrome interval gene), M97252; chicken KAL, L12144; a *Caenorhabditis elegans* gene homologous to KAL, Z81561; human HE4, X63187; rabbit epididymal protein BE-20, U26725; dog epididymal secretory protein CE4, S77395; a mouse EST clone similar to human HE4, AA276500; and human vasopressin, M25647. The numbers for interior branches refer to the bootstrap values for 100 replications. Bootstrap values of less than 40% are not given. The scale at the bottom is units of nucleotide substitutions per site. B: A possible scenario explaining the origin and development of trappin genes. Similarities in the exon-intron organizations and sequences among the REST, SLPI, and trappin genes suggest that the trappin genes originated from the ancestral REST gene through gene duplications and exon or domain shuffling of the second WAP motif of the SLPI gene, and are now members of the REST gene superfamily. Exons originating from REST genes are indicated by open boxes, and those originating from a SLPI gene are indicated by filled boxes. C: Co-localization of the trappin, REST, SLPI, HE4, and transglutaminase (TGM) genes on human chromosome 20 q12-13.1 (*31, 32, 35*).

and highly homologous introns that flank exon 2 of the genes (Fig. 5 and Ref. *14*). Similar cases have been reported for genes that become multimerized through recent duplications, such as genes of the murine serpins (Spi-2) (*42*) and snake venom phospholipase A2 isozymes (*43, 44*). According to the theoretical work of Ohta (*45, 46*) on the evolution of multigene families, a certain strength of adaptive selection is necessary for a useful gene family. Our previous data on trappin genes (*14*) and the results of studies on murine serpin and snake venom phospholipase A2 genes (*42-44*) demonstrated many more amino acid substitutions of these multigenes than synonymous changes especially around the active sites, as seen in many other multigene families that developed under positive Darwinian selection (*45, 46*). It is conceivable that the trappin multigenes facilitate the survival of the Suidae against invasive agents. The great diversity, observed among mammalian species, in the number and composition of the trappin genes can also be explained by Ohta's theory, that postulates that gene numbers can differ greatly among lineages even in the same environment.

The highly homologous sequences of introns of the trappin genes can be explained as a result of gene conversion, and there may still be potential for further conversion of the genes or gene locus that changes the number and composition of trappin genes. Our recent study on the trappin genes involving site-by-site analysis of shared nucleotide changes among trappin genes revealed that gene conversion occurred in the regions of introns 1 and 2 before and after the divergence of pig and wart hog (Kato, A., Furutani, Y., Yasue, H., and Hirose, S., manuscript in preparation). However, it remains unknown how only the introns have become homogenized, while the exons were allowed to maintain their variety. One explanation is that pigs with homogenized exons of trappin genes were eliminated through natural selection since the capacity to produce trappins with a variety of reactive site sequences must have been beneficial for the survival of the Suidae against invasive agents. It is also possible that some other unknown mechanisms exist that protect the variety of the reactive sites from homogenization.

We wish to thank Ichiro Tamechika for the helpful discussion, and Yohko Saruta, Setsuko Satoh and Kazuko Tanaka for the expert secretarial and technical assistance.

REFERENCES

1. Hirose, S., Furukawa, M., Tamechika, I., Itakura, M., Kato, A., Suzuki, Y., Kuroki, J., and Tachibana, S. (1996) Discovery of a new type of proteinase inhibitor family whose members have an anchoring sequence. *Adv. Exp. Med. Biol.* **389**, 43-49
2. Nara, K., Ito, S., Ito, T., Suzuki, Y., Ghoneim, M.A., Tachibana, S., and Hirose, S. (1994) Elastase inhibitor elafin is a new type of proteinase inhibitor which has a transglutaminase-mediated anchoring sequence termed "cementoin." *J. Biochem.* **115**, 441-448
3. Zeeuwen, P.L., Hendriks, W., de Jong, W.W., and Schalkwijk, J. (1997) Identification and sequence analysis of two new members of the SKALP/elafin and SPAI-2 gene family. Biochemical properties of the transglutaminase substrate motif and suggestions for a new nomenclature. *J. Biol. Chem.* **272**, 20471-20478
4. Molhuizen, H.O., Alkemade, H.A., Zeeuwen, P.L., de Jongh, G.J., Wieringa, B., and Schalkwijk, J. (1993) SKALP/elafin: an elastase inhibitor from cultured human keratinocytes. Purification, cDNA sequence, and evidence for transglutaminase cross-linking. *J. Biol. Chem.* **268**, 12028-12032
5. Steinert, P.M. and Marekov, L.N. (1995) The proteins elafin, filaggrin, keratin intermediate filaments, loricrin, and small proline-rich proteins 1 and 2 are isodipeptide cross-linked components of the human epidermal cornified cell envelope. *J. Biol. Chem.* **270**, 17702-17711
6. Kuroki, J., Hosoya, T., Itakura, M., Hirose, S., Tamechika, I., Yoshimoto, T., Ghoneim, M.A., Nara, K., Kato, A., Suzuki, Y., Furukawa, M., and Tachibana, S. (1995) Cloning, characterization, and tissue distribution of porcine SPAI, a protein with a transglutaminase substrate domain and the WAP motif. *J. Biol. Chem.* **270**, 22428-22433
7. Molhuizen, H.O. and Schalkwijk, J. (1995) Structural, biochemical, and cell biological aspects of the serine proteinase inhibitor SKALP/elafin/ESI. *Biol. Chem. Hoppe Seyler* **376**, 1-7
8. Hagstrom, J.E., Fautsch, M.P., Perdok, M., Vrabel, A., and Wieben, E.D. (1996) Exons lost and found. Unusual evolution of a seminal vesicle transglutaminase substrate. *J. Biol. Chem.* **271**, 21114-21119
9. Lilja, H. and Lundwall, A. (1992) Molecular cloning of epididymal and seminal vesicular transcripts encoding a semenogelin-related protein. *Proc. Natl. Acad. Sci. USA* **89**, 4559-4563
10. Lilja, H., Abrahamsson, P.A., and Lundwall, A. (1989) Semenogelin, the predominant protein in human semen. Primary structure and identification of closely related proteins in the male accessory sex glands and on the spermatozoa. *J. Biol. Chem.* **264**, 1894-1900
11. Wiedow, O., Schroder, J.M., Gregory, H., Young, J.A., and Christophers, E. (1990) Elafin: an elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence. *J. Biol. Chem.* **265**, 14791-14795
12. Furukawa, M., Suzuki, Y., Ghoneim, M.A., Tachibana, S., and Hirose, S. (1996) Cryptic origin of SPAI, a plasma protein with a transglutaminase substrate domain and the WAP motif, reveal-

ed by in situ hybridization and immunohistochemistry. *J. Biol. Chem.* **271**, 29517-21520
13. Ishizuka, N., Fukushima, Y., Urayama, O., and Akera, T. (1991) Na^+,K^+-ATPase inhibition by an endogenous peptide, SPAI-1, isolated from porcine duodenum. *Biochim. Biophys. Acta* **1069**, 259-266
14. Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A., Tachibana, S., and Hirose, S. (1996) Accelerated evolution in inhibitor domains of porcine elafin family members. *J. Biol. Chem.* **271**, 7012-7018
15. Schalkwijk, J., van Vlijmen, I.M., Alkemade, J.A., and de Jongh, G.J. (1993) Immunohistochemical localization of SKALP/elafin in psoriatic epidermis. *J. Invest. Dermatol.* **100**, 390-393
16. Saheki, T., Ito, F., Hagiwara, H., Saito, Y., Kuroki, J., Tachibana, S., and Hirose, S. (1992) Primary structure of the human elafin precursor preproelafin deduced from the nucleotide sequence of its gene and the presence of unique repetitive sequences in the prosegment. *Biochem. Biophys. Res. Commun.* **185**, 240-245
17. Lundwall, A. (1996) The cloning of a rapidly evolving seminal-vesicle-transcribed gene encoding the major clot-forming protein of mouse semen. *Eur. J. Biochem.* **235**, 424-430
18. Stetler, G., Brewer, M.T., and Thompson, R.C. (1986) Isolation and sequence of a human gene encoding a potent inhibitor of leukocyte proteases. *Nucleic Acids Res.* **14**, 7883-7896
19. Yasue, H. and Wada, Y. (1996) A swine SINE (PRE-1 sequence) distribution in swine-related animal species and its phylogenetic analysis in swine genome. *Anim. Genet.* **27**, 95-98
20. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680
21. Felsenstein, J. (1988) Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.* **22**, 521-565
22. Green, E.D. and Olson, M.V. (1990) Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **87**, 1213-1217
23. Rohrer, G.A., Alexander, L.J., Hu, Z., Smith, T.P., Keele, J.W., and Beattie, C.W. (1996) A comprehensive map of the porcine genome. *Genome Res.* **6**, 371-391
24. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425
25. Mulando, S.S., Harumi, T., Wada, Y., and Yasue, H. (1998) Localization of swine PRE-1 homologues in 13 loci of *Phacochoerus aethiopicus* and tyassu tajacu genomes, and their sequence divergence. *Anim. Genet.* **28**, 210-215
26. Strout, M.P., Marcucci, G., Bloomfield, C.D., and Caligiuri, M.A. (1998) The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia. *Proc. Natl. Acad. Sci. USA* **95**, 2390-2395
27. Grutter, M.G., Fendrich, G., Huber, R., and Bode, W. (1988) The 2.5 Å X-ray crystal structure of the acid-stable proteinase inhibitor from human mucous secretions analysed in its complex with bovine alpha-chymotrypsin. *EMBO J.* **7**, 345-351
28. Tsunemi, M., Matsuura, Y., Sakakibara, S., and Katsube, Y. (1996) Crystal structure of an elastase-specific inhibitor elafin complexed with porcine pancreatic elastase determined at 1.9 Å resolution. *Biochemistry* **35**, 11570-11576
29. Coronel, C.E., San Agustin, J., and Lardy, H.A. (1990) Purification and structure of caltrin-like proteins from seminal vesicle of the guinea pig. *J. Biol. Chem.* **265**, 6854-6859
30. Araki, K., Kuroki, J., Ito, O., Kuwada, M., and Tachibana, S. (1989) Novel peptide inhibitor (SPAI) of Na^+,K^+-ATPase from porcine intestine. *Biochem. Biophys. Res. Commun.* **164**, 496-502
31. Molhuizen, H.O., Zeeuwen, P.L., Olde Weghuis, D., Geurts van Kessel, A., and Schalkwijk, J. (1994) Assignment of the human gene encoding the epidermal serine proteinase inhibitor SKALP (PI3) to chromosome region 20q12→q13. *Cytogenet. Cell Genet.* **66**, 129-131
32. Schuler, G.D., Boguski, M.S., Stewart, E.A., Stein, L.D., Gyapay, G., Rice, K., White, R.E., Rodriguez-Tome, P., Aggarwal, A., Bajorek, E., Bentolila, S., Birren, B.B., Butler, A., Castle, A.B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P.J., Dibling, T., Drouot, N., Dunham, I., Duprat, S., East, C., Edwards, C., Fan, J.B., Fang, N., Fizames, C., Garrett, C., Green, L., Hadley, D., Harris, M., Harrison, P., Brady, S., Hicks, A., Holloway, E., Hui, L., Hussain, S., Louis-Dit-Sully, C., Ma, J., MacGilvery, A., Mader, C., Maratukulam, A., Matise, T.C., McKusick, K.B., Morissette, J., Mungall, A., Muselet, D., Nusbaum, H.C., Page, D.C., Peck, A., Perkins, S., Piercy, M., Qin, F., Quackenbush, J., Ranby, S., Reif, T., Rozen, S., Sanders, C., She, X., Silva, J., Slonim, D.K., Soderlund, C., Sun, W.L., Tabar, P., Thangarajah, T., Vega-Czarny, N., Vollrath, D., Voyticky, S., Wilmer, T., Wu, X., Adams, M.D., Auffray, C., Walter, N.A.R., Brandon, R., Dehejia, A., Goodfellow, P.N., Houlgatte, R., Hudson Jr, J.R., Ide, S.E., Iorio, K.R., Lee, W.Y., Seki, N., Nagase, T., Ishikawa, K., Nomura, N., Phillips, C., Polymeropoulos, M.H., Sandusky, M., Schmitt, K., Berry, R., Swanson, K., Torres, R., Venter, J.C., Sikela, J.M., Beckmann, J.S., Weissenbach, J., Myers, R.M., Cox, D.R., James, M.R., Bentley, D., Deloukas, P., Lander, E.S., and Hudson, T.J. (1996) A gene map of the human genome. *Science* **274**, 540-546
33. Lundwall, A. and Ulvsbäck, M. (1996) The gene of the protease inhibitor SKALP/elafin is a member of the REST gene family. *Biochem. Biophys. Res. Commun.* **221**, 323-327
34. Lundwall, A. and Lazure, C. (1995) A novel gene family encoding proteins with highly differing structure because of a rapidly evolving exon. *FEBS Lett.* **374**, 53-56
35. Ulvsbäck, M., Lazure, C., Lilja, H., Spurr, N.K., Rao, V.V., Loffler, C., Hansmann, I., and Lundwall, A. (1992) Gene structure of semenogelin I and II. The predominant proteins in human semen are encoded by two homologous genes on chromosome 20. *J. Biol. Chem.* **267**, 18080-18084
36. Harris, S.E., Harris, M.A., Johnson, C.M., Bean, M.F., Dodd, J.G., Matusik, R.J., Carr, S.A., and Crabb, J.W. (1990) Structural characterization of the rat seminal vesicle secretion II protein and gene. *J. Biol. Chem.* **265**, 9896-9903
37. Bingle, L. and Tetley, T.D. (1996) Secretory leukoprotease inhibitor: partnering alpha 1-proteinase inhibitor to combat pulmonary inflammation. *Thorax* **51**, 1273-1274
38. Fritz, H. (1988) Human mucus proteinase inhibitor (human MPI). Human seminal inhibitor I (HUSI-I), antileukoprotease (ALP), secretory leukocyte protease inhibitor (SLPI). *Biol. Chem. Hoppe Seyler* **369**, 79-82
39. Rosato, E., Peixoto, A.A., Gallippi, A., Kyriacou, C.P., and Costa, R. (1996) Mutational mechanisms, phylogeny, and evolution of a repetitive region within a clock gene of *Drosophila melanogaster*. *J. Mol. Evol.* **42**, 392-408
40. Kim, H.S., Lyons, K.M., Saitoh, E., Azen, E.A., Smithies, O., and Maeda, N. (1993) The structure and evolution of the human salivary proline-rich protein gene family. *Mamm. Genome* **4**, 3-14
41. Rodakis, G.C., Lecanidou, R., and Eickbush, T.H. (1984) Diversity in a chorion multigene family created by tandem duplications and a putative gene-conversion event. *J. Mol. Evol.* **20**, 265-273
42. Inglis, J.D. and Hill, R.E. (1991) The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain. *EMBO J.* **10**, 255-261
43. Nakashima, K., Ogawa, T., Oda, N., Hattori, M., Sakaki, Y., Kihara, H., and Ohno, M. (1993) Accelerated evolution of *Trimeresurus flavoviridis* venom gland phospholipase A2 isozymes. *Proc. Natl. Acad. Sci. USA* **90**, 5964-5968
44. Nakashima, K., Nobuhisa, I., Deshimaru, M., Nakai, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y., Hattori, S., and Ohno, M. (1995) Accelerated evolution in the protein-coding regions is universal in crotalinae snake venom gland phospholipase A2 isozyme genes. *Proc. Natl. Acad. Sci. USA* **92**, 5605-5609
45. Ohta, T. (1991) Multigene families and the evolution of complexity. *J. Mol. Evol.* **33**, 34-41
46. Ohta, T. (1988) Multigene and supergene families in *Oxford Surveys in Evolutionary Biology* V (Harvey, P.H. and Partridge, L., eds.) pp. 41-65, Oxford University Press, Oxford
47. Shimamura, M., Yasue, H., Ohshima, K., Abe, H., Kato, H., Kishiro, T., Goto, M., Munechika, I., and Okada, N. (1997) Molecular evidence from retroposons that whales form a clade within even-toed ungulates. *Nature* **388**, 666-670