Plant Molecular Biology **33:** 187–192, 1997.

Short communication

Aspartic proteinase genes in the Brassicaceae *Arabidopsis thaliana* and *Brassica napus*

Kathleen D'Hondt[1,2,5], Sylvia Stack[3], Steve Gutteridge[3], Joel Vandekerckhove[2], Enno Krebbers[1,3,+] and Susannah Gal[4,*]

[1]*Plant Genetic Systems, J. Plateaustraat 22, B-9000 Gent, Belgium;* [2]*Laboratorium Fysiologische Chemie, Universiteit Gent, K. Ledeganckstraat 35, B-9000 Gent, Belgium;* [3]*DuPont Agricultural Products, Experimental Station 402 Wilmington, Delaware 19880-0402, USA (*[+]*present address);* [4]*Department of Biological Sciences, The State University of New York at Binghamton, Binghamton, NY 13902–6000, USA (*[*]*author for correspondence);* [5]*present address: Laboratory for Monoclonal Antibodies, P.O. Box 8123, 6700 ES Wageningen, Netherlands*

Received 27 March 1996; accepted in revised form 1 October 1996

Key words: aspartic proteinase, gene isolation, intron/exon arrangement, protein processing, pulmonary surfactant-associated protein, saposin C

Abstract

Active aspartic proteinase is isolated from *Brassica napus* seeds and the peptide sequence is used to generate primers for PCR. We present here cDNA and genomic clones for aspartic proteinases from the closely related Brassicaceae *Arabidopsis thaliana* and *Brassica napus*. The *Arabidopsis* cDNA represents a single gene, while *Brassica* has at least 4 genes. Like other plant aspartic proteases, the two Brassicaceae enzymes contain an extra protein domain of about 100 amino acids relative to the mammalian forms. The intron/exon arrangement in the *Brassica* genomic clone is significantly different from that in mammalian genes. As the proteinase is isolated from seeds, the same tissue where 2S albumins are processed, this implies expression of one of the aspartic proteinase genes there.

Aspartic proteinases appear to be the simplest subclass of peptidases known, containing only three families, and are found in numerous organisms including viruses, fungi, animals and plants [24]. A role for these enzymes in protein processing has been demonstrated in polyprotein processing of the AIDS virus [5], in enzyme precursor processing in yeast and animal cells [11, 15, 21], and with seed storage protein precursor processing *in vitro* in plants [8, 26].

Isolation of the *Arabidopsis* cDNA for an aspartic proteinase was facilitated by the availability of several expressed sequence tags (ESTs) with significant homology to the barley aspartic proteinase cDNA [20]. Three such ESTs were obtained from the *Arabidopsis* Biological Resource Center (ABRC, Columbus, OH) which represent homology to different parts of the aspartic proteinase sequence. These ESTs were shown to be homologous on Southern blots (data not shown), and the longest was used to screen a cDNA library [16]. A 1.7 kb clone was isolated (pSG2J2) and fully sequenced. The deduced protein sequence confirmed the high homology to known plant aspartic proteinases, but indicated the clone did not code for approximately 25 amino acids representing the amino terminal signal peptide (Fig. 1).

The previously characterized *Brassica* aspartic proteinase [8] was further purified to obtain partial amino acid sequence. As detailed in the legends, a critic acid extract of *B. napus* seed cake was concentrated, subjected to pepstatin A affinity chromatography, and the resulting eluate further separated using size exclusion chromatography. Coomassie staining of a polyacryl-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers U51306 (pSG2J2), U55033 (pKD421) and U55032 (pKD425).

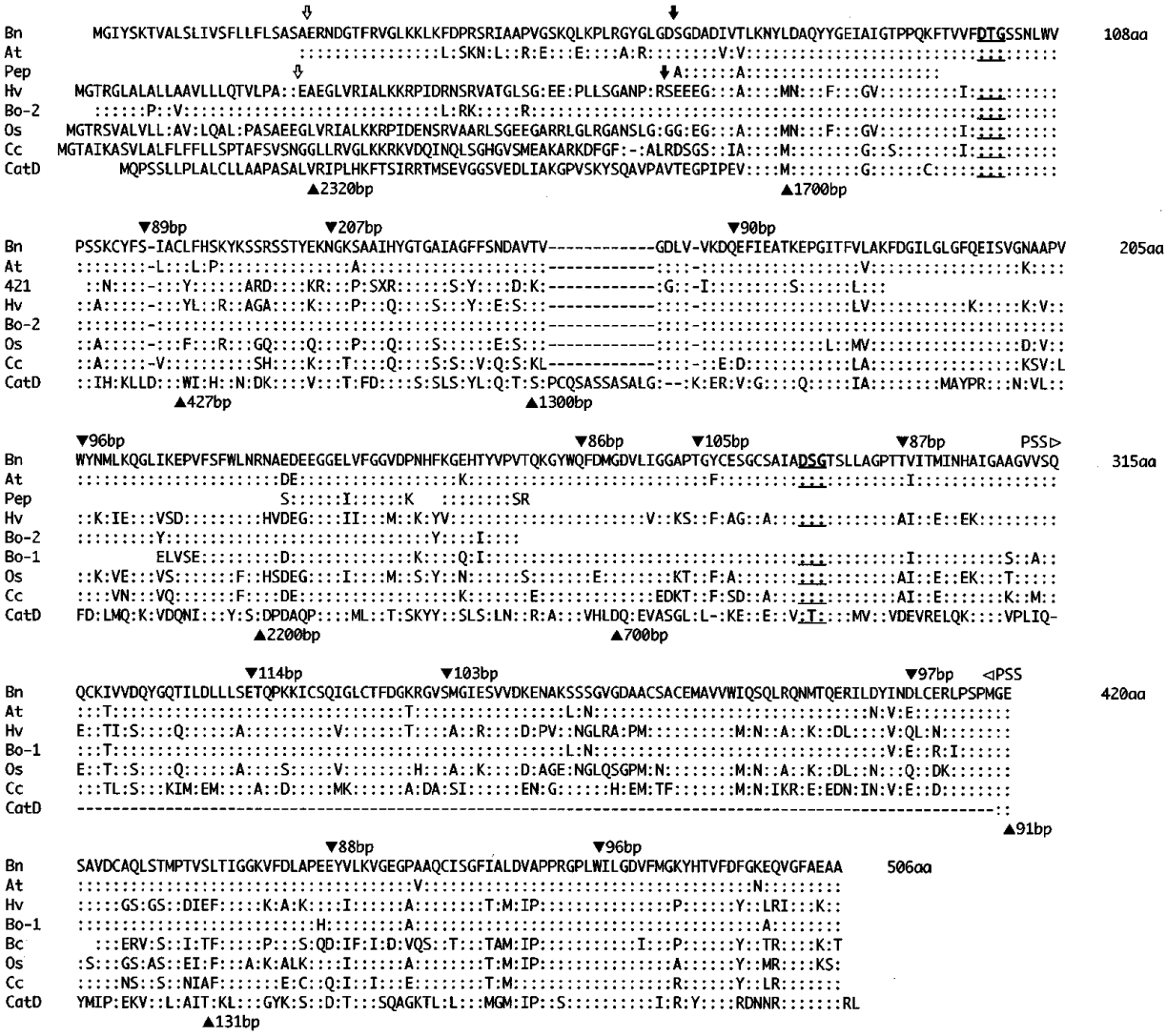

Figure 1.

amide gel demonstrated that two major protein bands co-eluted with a proteinase activity that was inhibited by pepstatin A (Fig. 2). The specificity of the proteinase activity isolated was confirmed using a synthetic peptide corresponding to the internal processed fragment of one of the *Arabidopsis* 2S albumin isoforms previously shown to be cleaved by an aspartic proteinase ([8]; data not shown). To obtain amino acid sequence, these proteins were separated on a preparative gel, blotted onto a Problott membrane and visualized with amido black. The two protein bands (28 kDa and 35 kDa) were excised, and sequence obtained from amino-terminal or tryptic peptides. The resulting sequences showed that the 28 kDa peptide had significant homology to the amino acid sequences of known aspartic proteinases (Fig. 1) while the 35 kDa band showed no homology to any known sequence (data not shown).

PCR primers were designed for these regions as well as for regions containing highly conserved amino acids around the two aspartic acid residues in the active site in previously characterized aspartic proteinases. Using *Brassica* genomic DNA as a template, two PCR amplification fragments were cloned and sequenced (the insert size of pKD420 was 670 bp and for pKD421 was 570 bp). The sequences were similar

but not identical (see Fig. 1). Genomic clones encoding the complete *B. napus* aspartic proteinase gene were isolated from an enriched library using pKD420 as a probe (see Fig. 1). One clone (pKD425) contained a 5.6 kb insert and was completely sequenced. Intron-exon boundaries were determined on the basis of homology with the *Arabidopsis* cDNA and concensus sequences for splice junctions. As expected, the deduced amino acid sequence shows significant homology of the *Arabidopsis* and *Brassica* enzymes beyond the putative signal sequence (marked in Fig. 1), including the proposed active site aspartic acids and conserved cysteine residues. The sequence of the genomic clone (pKD425) is identical to pKD420 in the region of overlap, but different from the other PCR fragment (pKD421) and from the peptide sequences (Fig. 1), which suggested that there are at least two genes for aspartic proteinases in *B. napus*.

Southern blotting demonstrated that the aspartic proteinases are coded in *B. napus* by at least 4 genes but only a single gene in *Arabidopsis*. Genomic DNA was restricted using a number of different enzymes, separated on a gel, blotted and probed with either pSG2J2 for the *Arabidopsis* samples and pKD420 and pKD421 for the *Brassica* samples. The results indicate that the *Arabidopsis* gene is likely to be single copy as single bands were detected in several digests (Fig. 3A). The aspartic proteinases in *B. napus* appear to be encoded by a small gene family (Fig. 3B). For instance, digestion of genomic DNA with *Bgl*II shows 3 major bands with the pKD420 as probe, but two distinct bands when pKD421 was used as probe (Fig. 3B). Neither probe contains a restriction site for any of the enzymes used. Furthermore, the two probes do not cross hybridize at the stringency used in this experiment (Fig. 3B; lanes 420 and 421). Taken together, the banding pattern in the Southern blots, and the differences between the sequences from the PCR and genomic clones, as well as the peptides, the data are consistent with at least 4 genes being present in *B. napus*. The purification of active protein from seeds of *B. napus* (this work) and *A. thaliana* (A. Mutlu and S. Gal, manuscript in preparation) demonstrates that these genes are expressed in seeds.

Comparison of all of the plant aspartic proteinases sequenced to date shows limited homology in the putative signal peptide and amino-terminal propeptide, but clear homology in the downstream regions (Fig. 1).

Figure 1. Comparison of aspartic protease sequences. Sequences were deduced from the coding sequences of the *Brassica napus* genomic clone pKD425 (Bn), the *Arabidopsis thaliana* cDNA pSG2J2 (At), the partial *B. napus* genomic PCR fragment pKD421 (421) and, partial peptide sequences of a purified *B. napus* aspartic protease (see Fig. 2) (Pep). These sequences are compared to previously published sequences from barley (accession number X56136) (Hv) [27], *Cynara cardunculus* (accession number X81984) (Cc) [6], *Brassica oleracea* (accession number X80067) (Bo-2) [10], rice (accession number D32144) (Os) [1] and human cathepsin D (accession number M11233) [9], as well as to partial sequences in GenBank from *Brassica campestris* (accession number L33639) (Bc), and *B. oleracea* (accession number X77260) (Bo-1). The 'X' in the 421 sequence is due to a sequencing ambiguity but the codon would form either an I or M at that position. The numbering is based on the Bn sequence and a colon ':' indicates the sequence is identical to the Bn sequence on top. Gaps are included to maximize the alignment. The solid arrow indicates the end of the signal peptide determined biochemically for the barley clone [26] while that predicted for the *Brassica* gene is at a similar location [23, 35]. The open arrow marks the beginning of the mature protein based on amino-terminal sequences of the barley protein [31] and the sequenced peptides from the *Brassica* enzyme. PSS indicates the region unique to plant genes. The positions and sizes of the introns in the *B. napus* and human genes are indicated as filled arrowheads with numbers above or below the sequence. pKD421 contains the first 3 introns found in pKD425 at the same positions, but the sizes (104, 88 and 91 bp, respectively) and sequences are not identical. Complete DNA sequences, including introns, for pKD421, pKD425 and pSG2J2 are to be found under the accession number U55033, U55032, and U51036. DNA manipulations were done according to established procedures [28]. The *Arabidopsis* cDNA was obtained using EST T21482 (homologous to ESTs T20862 and T20903) to probe a cDNA library from the ABRC (CD4-15) which was made from seedling hypocotyls from *A. thaliana* ecotype Columbia. About 30 000 plaques were screened to obtain a positive clone, pSG2J2. The clone was completely sequenced manually using the method of Sanger [29]. Gels similar to that shown in Fig. 2 were run and blotted on a Pro Blott (Applied Biosystems) membrane. The membrane was stained with amido black, the 28 kDa band was excised and treated with trypsin (except for the sample used for amino-terminal sequencing). The tryptic peptides were isolated and sequenced [2]. pKD420 and pKD421 (see text) were isolated using the PCR primers PCR2 GAIACIGGII SIISIAAIYT ITGGGTICC and PCR3 CCIAIICCIA RIATICCRTC RAA (I = inosine). Amplification was done using *Taq* polymerase (Perkin Elmer) with 100 ng genomic *B. napus* DNA as template, 100 pmol of each primer and an annealing temperature of 40 °C. The two bands resulting were further amplified individually with the band stab method [3]. The amplification fragments were isolated from gel with Qiaex (Qiagen), treated with T4 polymerase and T4 kinase, and cloned into pGEM 3Zf(+) (Promega) restricted with *Sma*I. Recombinant clones were detected with blue-white screening. The inserts were sequenced using an ABI automated sequencer. pKD425 was isolated as follows; 150 μg *B. napus* DNA was digested overnight with *Eco*RI and separated overnight on a preparative 1% low-melting agarose gel. The gel was sliced into pieces of 0.5 cm in the size range between 3 kb and 11 kb. The DNA was extracted from the gel matrix with phenol and precipitated. The different samples were then assayed using a Southern blot. The fractions for which positive signals were detected were used for cloning in Lambda arms cut with *Eco*RI and dephosphorylated (Lambda ZAPR II/*Eco*RI/CIAP, Stratagene). The constructs were then packaged using GigapackR II packaging extract (Stratagene); 5×10^5 pfu were screened with pKD420. Single colonies were isolated and inserts subcloned using *in vivo* excision according to the manufacturer's instructions (Stratagene), such that the fragment was cloned in pBluescript SK. After preliminary analysis, pKD425 was selected for complete sequencing using an ABI automated sequencer.

Striking similarities in the sequence coincide with the active site aspartic acid residues and the alignment of the conserved cysteine residues which are common with the mammalian aspartic proteinases. The *Brassica* genomic sequence (pKD425) is clearly the homologue of one of the *B. oleracea* sequences (Bo-2 in Fig. 1), as might be expected since the tetraploid *B. napus* is the result of a cross between *B. oleracea* and *B. campestris*. The other partial *B. napus* sequences (nucleic acid and protein) do not have clear counterparts in the partial sequences available for the other *B. oleracea* (Bo-1) and *B. campestris* (Bc) isoforms, but it is likely that the multiple *B. napus* genes will have homologues in one or the other of the progenitor species.

The plant aspartic proteinases have diverged from the mammalian forms of these enzymes in both the protein coding region and the intron/exon arrangement. Relative to the mammalian sequences, plant aspartic proteases all contain an insertion of about 100 amino acids called the plant specific sequence (PSS in Fig. 1). A blast search using this region alone shows slight homology to saposin C and pulmonary surfactant-associated proteins with a conserved arrangement of the cysteine residues [12, 17, 34]. This might suggest an interaction of the enzyme with lipids via this exposed domain. Recent work with human cathepsin D [36] has shown a transient association of the enzyme during synthesis with a protein having antigenic similarity to saposin C. These researchers propose that this interaction may facilitate the lysosomal targeting of the enzyme which would be consistent with work studying another enzyme [33]. The possible role of this domain awaits parallel *in vitro* studies with mammalian and plant proteinases.

All of the plant splice junctions met the canonical GT/AG requirement, although immediately adjacent sequences were less conserved (sequences available in the databank). As shown in Fig. 1, the position and size of the 12 introns in the *Brassica* genomic clone do not match those found in the human cathepsin D gene [25]. The arrangement in cathepsin D is representative of that in several other mammalian aspartic proteinase genes [13, 14, 19, 22, 32]. With the exception of the partial PCR clone (pKD421, Fig. 1), which has three introns in analogous positions but which differ in sequence and length from those of pKD425, no other plant genomic sequences are available. Thus we are unable to determine if this intron arrangement is a general feature of plant aspartic proteinases. One mammalian intron is located near the site of the PSS, suggesting the possibility that an alternative splicing event resulted in the incorporation of the PSS in the plant messages. The length of this intron is quite variable in the different mammalian genes some containing less sequence than would be necessary to code for the entire PSS [13, 14, 19, 22, 25, 32]. Only a few examples of genes for other proteins common to both

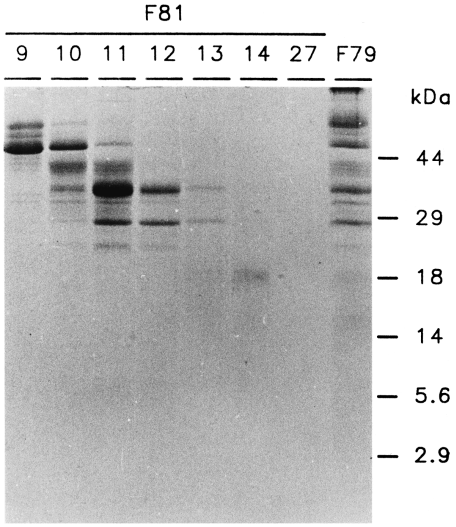

Figure 2. Polyacrylamide gel analysis of gel filtration fractions of *Brassica napus* seed extracts. About 1/20th of the fractions was loaded on a 15% polyacrylamide gel, which was stained with Coomassie Blue. The numbers on top of the lanes refer to the fractions collected after gel filtration. F81 is number of the gel filtration run. F79 of 1/80 of the fraction obtained after elution from the pepstatin A column. The molecular size markers are indicated in kDa. 100 g seed cake of *B. napus* was extracted in 500 ml 100 mM citric acid buffer pH 3.5, 5 mM PMSF, 5 mM EDTA for 2 h at 4 °C. The extract was isolated by centrifugation at 5000 × g. The supernate was clarified by centrifugation in an SS-34 rotor at 13 000 rpm and subsequently concentrated by ultrafiltration through a PM-30 filter to a final volume of 50 ml. The pH was adjusted to 3.5 with 2 M unbuffered citrate solution and the salt concentration to 200 mM NaCl in a total volume of 100 ml. After centrifugation this sample was loaded onto a 5 ml pepstatin A agarose column (Sigma) and eluted as described [8]. A fraction of about 20 ml corresponding to a peak detected at 280 nm was collected and lyophilized after dialysis against deionized water. The sample was resuspended in 250 μl water, and 200 μl of this sample was loaded on a Superdex 75 HR 10/30 gel filtration column (Pharmacia). The column was run in 150 mM NH_4HCO_3. Fractions of 0.5 ml were collected, lyophilized and resuspended in 50 μl deionized water. For each fraction, enzymatic activity and its pepstatin sensitivity were analyzed as described ([31] and references therein), using bovine hemoglobin as a substrate. The enzyme generates trichloroacetic acid (TCA) soluble peptides at pH 3.7 and 30 °C from the hemoglobin substrate. The concentration of the TCA soluble peptides is a measure for the protease activity. One unit corresponds to an enzyme activity that generates TCA soluble products in 1 h at 30 °C equivalent to 1 mg BSA. The specific activities in the different fractions were as follows: F9 5.2U/mg protein, F10 19.6, F11 39.7, F12 61.3, and F13 31.6.

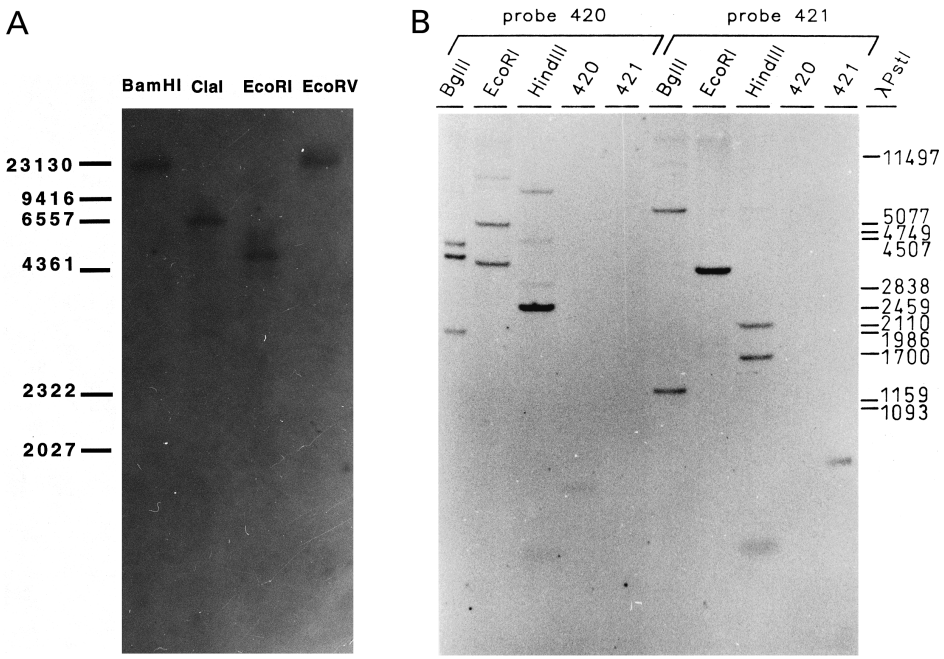

Figure 3. Southern blot analysis of *Arabidopsis* and *Brassica* genomic DNA. A. *Arabidopsis* genomic DNA isolated from RLD ecotype via the method of Dellaporta *et al.* [7] (10 μg) was restricted with the enzymes indicated and probed with the digoxigenin labeled insert from pSG2J2 (Boehringer Mannheim) in 50% formamide, 5 $\times$ SSC, 0.1% sarkosyl, 0.02% SDS and 2% blocking agent (Boehringer Mannheim) at 42 °C overnight. The blot was washed in 2 $\times$ SSC 0.1% SDS and 0.5 $\times$ SSC 0.1% SDS at 42 °C and then processed for detection of the digoxigenin (Boehringer Mannheim). B. *Brassica napus* genomic DNA was restricted with the enzymes indicated. The lanes marked 420 and 421 contain 15 pmol of plasmids pKD420 and pKD421, respectively. The left side of the blot was probed with pKD420, the right with pKD421. The final wash was 5 min in 0.1 $\times$ SSC, 0.1% SDS at 65 °C. The positions of the molecular weight markers are indicated and the sizes given in bp.

animals and plants have been isolated [4, 18, 30] and in many cases the intron/exon arrangement is relatively conserved between these two divergent species. Our data indicate substantial addition and/or removal of introns in the aspartic proteinase genes since the evolutionary divergence of the two kingdoms making it significantly different from other common genes thus far isolated.

While this manuscript was under review, a rice genomic sequence for the seed aspartic proteinase was submitted to the database (Acc. No. D32165). The intron/exon placement is identical between these sequences although 3 of the introns in the rice are significantly larger than those in *B. napus*. This indicates that placement of introns in the plant aspartic proteinase genes predates the divergence of monocot and dicot species.

Acknowledgements

We thank Jef Seurinck for sequencing pKD420 and pKD421, Rikus Pomp for help in the early stages of sequencing pKD425, and Jose Van Damme for peptide sequencing. We would also like to thank Sandeep Gulati for help in isolating pSG2J2 and Jane Tien for help with some of the sequencing of this plasmid. S. Gal was supported in part by a Binghamton University Faculty Summer Research Support grant, while K.D. was supported by a grant from the Belgian IWT.

References

1. Asakura T, Watanabe H, Abe K, Arai S: Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases. Eur J Biochem 232: 77–83 (1995).
2. Bauw G, Rasmussen HH, Van Den Bulcke M, Van Damme J, Puype M, Gesser B, Celes JE, Vandekerckhove J: 2-Dimensional gel-electrophoresis, protein electroblotting and

microsequencing: a direct link between protein and genes. Electrophoresis 11: 528–536 (1990).
3. Bjourson AJ, Cooper JE: Band-stab PCR: a simple technique for the purification of individual PCR products. Nucl Acids Res 20: 4675 (1992).
4. Claes V, Kettmann R, Burny A: Structure of the gene encoding pig phosphoglucose isomerase. Gene 150: 235–241 (1994).
5. Co E, Koelsch G, Lin Y, Ido E, Hartsuck JA, Tang J: Proteolytic processing mechanisms of a miniprecursor of the aspartic protease of human immunodeficiency virus type I. Biochem 33: 1248–1254 (1994).
6. Cordeiro MC, Xue Z-T, Pietrzak M, Pais MS, Brodelius PE: Isolation and characterization of a cDNA from flowers of *Cynara cardunculus* encoding cyprosin (an aspartic proteinase) and its use to study the organ-specific expression of cyprosin. Plant Mol Biol 24: 733–741 (1994).
7. Dellaporta SL, Wood J, Hicks JB: A plant DNA minipreparation: version II. Plant Mol Biol Rep 1: 19–21 (1983).
8. D'Hondt K, Bosch D, Van Damme J, Goethals M, Vandekerckhove J, Krebbers E: An aspartic proteinase present in seeds cleaves *Arabidopsis* 2S albumin precursors *vitro*. J Biol Chem 268: 20884–20891 (1993).
9. Faust P, Kornfeld S, Chirgwin JM: Cloning and sequence analysis of cDNA for human cathepsin D. Proc Natl Acad Sci USA 82: 4910–4914 (1985).
10. Fujikura Y, Karssen CM: Molecular studies on osomoprimed seeds of cauliflower: a partial amino acid sequence of a vigour-related protein and osmopriming-enhanced expression of putative aspartic protease. Seed Sci Res 5: 177–181 (1995).
11. Gottschalk S, Waheed A, Schmidt B, Laidler P, von Figura K: Sequential processing of lysosomal acid phosphatase by a cytoplasmic thiol proteinase and a lysosomal aspartyl proteinase. EMBO J 8: 3215–3219 (1989).
12. Guruprasad K, Tormakangas K, Kervinen J, Blundell TL: Comparative modelling of barley-grain aspartic proteinase: a structural rationale for observed hydrolytic specificity. FEBS Lett 352: 131–136 (1994).
13. Hidaka M, Sasaki K, Uozumi T, Beppu T: Cloning and structural analysis of the calf prochymosin gene. Gene 43: 197–203 (1986).
14. Hobart PM, Fogliano M, O'Connor BA, Schaefer IM, Chirgwin JM: Human renin gene: Structure and sequence analysis. Proc Natl Acad Sci USA 81: 5026–5030 (1984).
15. Jones EW: Three proteolytic systems in the yeast *Saccharomyces cerevisiae*. J Biol Chem 266: 7963–7966 (1991).
16. Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR: CTR1, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the Raf family of protein kinases. Cell 72: 427–441 (1993).
17. Kuroki Y, Voelker DR: Pulmonary surfactant proteins. J Biol Chem 269: 25943–25946 (1994).
18. Marchionni M, Gilbert W: The triosephosphate isomerase gene from maize: introns antedate the plant-animal divergence. Cell 46: 133–141 (1986).
19. Miyazaki H, Fukamizu A, Hirose S, Hayashi T, Hori H, Ohkubo H, Nakanishi S, Murakami K: Structure of the human renin gene. Proc Natl Acad Sci USA 81: 5999–6003 (1984).
20. Newman T, de Bruijn FJ, Green P, Keegstra K, Kende H, McIntosh L, Ohlrogge J, Raikhel N, Somerville S, Thomashow M, Retzel E, Somerville C: Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiol 106: 1241–1255 (1994).
21. Nishimura Y, Kawabata T, Furuno K, Kato K: Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. Arch Biochem Biophys 271: 400–406 (1989).
22. Oerd T, Kolmer M, Villems R, Saarma M: Structure of the human genomic region homologous to the bovine prochymosin-encoding gene. Gene 91: 241–246 (1990).
23. Perlman D, Halvorson H: A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J Mol Biol 167: 391–409 (1983).
24. Rawlings ND, Barrett AJ: Families of aspartic preptidases, and those of unknown catalytic mechanism. Meth Enzymol 248: 105–120 (1995).
25. Redecker B, Heckendorf B, Grosch H-M, Mersmann G, Hasilik A: Molecular organization of the human cathepsin D gene. DNA Cell Biol 10: 423–431 (1991).
26. Runeberg-Roos P, Kervinen J, Kovaleva V, Raikhel NV, Gal S: The aspartic proteinase of barley is a vacuolar enzyme that processes probarley lectin *in vitro*. Plant Physiol 105: 321–329 (1994).
27. Runeberg-Roos P, Törmäkangas K, Östman A: Primary structure of a barley-grain aspartic proteinase. A plant aspartic proteinase resembling mammalian cathepsin D. Eur J Biochem 202: 1021–1027 (1991).
28. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, NY (1989).
29. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
30. Sangwan V, Lenvik TR, Gant JS: The *Arabidopsis thaliana* ribosomal protein S15 (*rig*) gene. Biochim Biophys Acta 1216: 221–226 (1993).
31. Sarkkinen P, Kalkkinen N, Tilgmann C, Siuro J, Kervinen J, Mikola L: Aspartic proteinase from barley grains is related to mammalian lysosomal cathepsin D. Planta 186: 317–323 (1992).
32. Sogawa K, Fujii-Kuriyama Y, Mizukami Y, Ichihara Y, Takahasi K: Primary structure of human pepsinogen gene. J Biol Chem 258: 5306–5311 (1983).
33. Staab JF, Ginkel DL, Rosenberg GB, Munford RS: A saposin-like domain influences the intracellular localization, stability, and catalytic activity of human acyloxyacyl hydrolase. J Biol Chem 269: 23736–23742 (1994).
34. Vaccaro AM, Salvioli R, Barca A, Tatti M, Ciaffoni F, Maras B, Siciliano R, Zappacosta F, Amoresano A, Pucci P: Structural analysis of saposin C and B: complete localization of disulfide bridges. J Biol Chem 270: 9953–9960 (1995).
35. Von Heijne G: A new method for predicting signal sequence cleavage sites. Nucl Acids Res 14: 4683–4690 (1986).
36. Zhu Y, Conner GE: Intermolecular association of lysosomal protein precursors during biosynthesis. J Biol Chem 269: 3846–3851 (1994).