

RESEARCH PAPER

NtCP56, a new cysteine protease in *Nicotiana tabacum* L., involved in pollen grain development

Xiao-mei Zhang, Ying Wang, Xiao-meng Lv, Hui Li, Peng Sun, Hai Lu* and Feng-lan Li

College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing 100083, PR China

Received 3 November 2008; Revised 2 January 2009; Accepted 20 January 2009

Abstract

Proteinases play a critical role in developmental homeostasis and in response to environmental stimuli. Our present research reports that a new cysteine protease, NtCP56, is involved in the development of pollen grains in *Nicotiana tabacum* L. The NtCP56 gene, which encodes a protein of 361 amino acid residues with a calculated molecular mass of 40 kDa, is strongly expressed in anthers. The recombinant NtCP56 showed a high activity towards casein. Kinetic analysis revealed a K_m of 2.20 mg ml⁻¹ and V_{max} of 11.07 µg ml⁻¹ min⁻¹. The recombinant NtCP56 retained more than 50% of its maximum enzymatic activity from 20 °C to 60 °C with an optimum T_m range of 30–50 °C. The enzyme had a maximum activity at approximately pH 6.5. Suppression of the NtCP56 gene in anti-sense transgenic tobaccos resulted in the sterility of pollen grains. Our data indicated that, as a cysteine protease, NtCP56 might play an important role in pollen development.

Key words: Cysteine protease, NtCP56, pollen, tobacco.

Introduction

Male reproductive processes in flowering plants take place in the anther. In anthers, archesporial cells are thought to divide periclinally to give rise to parietal cells toward the exterior and to sporogenous cells toward the interior (Davis, 1966; Scott *et al.*, 2004). Parietal cells finally differentiate into inner anther walls, including the endothecium, the middle layer and the tapetum, and sporogenous cells undergo several rounds of mitotic division and differentiate to acquire a meiotic cell fate (Nonomura *et al.*, 2007).

Given the significance of male sterility to the agricultural industry, male gametophyte development has always been a major issue in plant studies and a few genes have been identified which control or participate in the development of pollen grains. In *Arabidopsis*, diverse factors including hormones, transcription factors, and receptors, have been reported to affect male gametophyte development (Sanders *et al.*, 1999; Scott *et al.*, 2004; Mizuno *et al.*, 2007). For instance, a receptor-like protein kinase 2 (RPK 2), is a factor controlling anther development. The RPK2 T-DNA, inserted mutants displayed male sterility due to defects in

anther dehiscence and pollen maturation (Mizuno *et al.*, 2007).

More genes participate in the development of pollen grains by controlling tapetum formation and degradation. Two highly similar *Arabidopsis* LRR (leucine-rich repeat) receptor-like protein kinases, SERK1 and SERK2 (somatic embryogenesis receptor kinase), were shown to be required for tapetum formation in a way similar to that of EMS1/EXS (excess microsporocytes1/extra sporogenous cell) (Albrecht *et al.*, 2005; Colcombet *et al.*, 2005; Li *et al.*, 2006). The *Arabidopsis* MS1 gene encodes a nuclear protein with a PHD-finger motif which has been shown to participate in the normal formation of the tapetum (Wilson *et al.*, 2001; Ito and Shinozaki, 2002). In rice, an undeveloped tapetum1 (*Udt1*) gene was also shown to be vital to tapetum differentiation and the formation of microspores (Jung *et al.*, 2005).

Until recently, the rice (*Oryza sativa*) tapetum degeneration retardation gene (TDR), encoding a putative basic helix–loop–helix protein, was found to be required in the degradation of the tapetum. In the TDR mutant rice,

* To whom correspondence should be addressed: luhai1974@yahoo.com
© 2009 The Author(s).

tapetum degeneration was retarded and the middle layer cells persisted, accompanied by aborted pollen development and complete male sterility. It was also verified that the direct targets of TDR were a cysteine protease OsCP1 and its inhibitor (Li *et al.*, 2006).

Cysteine proteases (CPs) have been reported to be the major enzymes responsible for the hydrolysis of most of the storage proteins, hordeins and glutelin, in crops (Lee *et al.*, 2004). CPs also contribute to the regulation of programmed cell death (Solomon *et al.*, 1999). For instance, the expression of *SmCP*, a *S. melongena* gene encoding a cysteine proteinase, coincides with several, rather than with specific, events in developmentally regulated programmed cell death in brinjal (Xu and Chye, 1999).

In this study, a new cysteine protease gene isolated from *Nicotiana tabacum* L. is reported. The *NtCP56* gene was strongly expressed in anthers. The recombinant NtCP56 protein showed activity towards casein proteolysis *in vitro*. In anti-sense NtCP56 transgenic tobacco anthers, the degradation of the tapetum was retarded and pollen development was aborted. These results provide important insights into the crucial role of *NtCP56* for pollen development.

Materials and methods

Molecular cloning

The GenBank database was searched using the BLAST program (<http://www.ncbi.nlm.nih.gov/BLAST/>) with *Nicotiana tabacum* cysteine protease cDNA (GenBank accession no. AY881010) as query sequence. One *Nicotiana tabacum* partial mRNA sequence (GenBank accession no. U57825) encoding an endopeptidase-like protein with high similarity to *Nicotiana tabacum* cysteine protease cDNA was identified. Based on this sequence, five primers (primer 1: 5'-TTGTATGAGAGATGGAGAAGCCAT-3'; primer 2: 5'-AGTTGAAAATGCCAGCAACTTCC-3'; primer 3: 5'-TGTATGCTCACGAGGATAG-3'; primer 4: 5'-CTGATCTAGAGGTACCGATCC-3'; and primer 5: 5'-TTTAAACATGAAGAAATTACTACTAC-3') were designed to amplify *Nicotiana tabacum NtCP56* cDNA.

Total RNA from anthers of *Nicotiana tabacum* L., isolated using the Trizol reagent, was used for a reverse-transcription polymerase chain reaction (RT-PCR). The first strand cDNA was synthesized in a volume of 25 μ l containing about 2 μ g RNA, 4 μ l reaction buffer, 1 μ l of 10 mM dNTP, 0.5 μ l RNase inhibitor, 1 μ l adaptor primer (5'-CTGATCTAGAGGTACCGATCCTTTTTTTTTT-TTTTTT-3'), and 1 μ l AMV reverse transcriptase. The transcription reaction was performed at 42 °C for 1 h and terminated at 65 °C for 10 min (Li *et al.*, 2002).

A specific cDNA fragment was first amplified by using primers 1 and 2. To isolate the complete region of this gene, a 3' race PCR was carried out by using primer 3 and primer 4 according to the instructions of the manufacturer (Takara 3' full race kit, Japan). The PCR was carried out as follows:

94 °C for 5 min, followed by 25 cycles of 94 °C for 1 min, 56 °C for 1 min, and 72 °C for 2 min. To verify the integrity of the cDNA, sequences of the gene RT-PCR were carried out to amplify the full-length cDNA using primers 4 and 5. The PCR thermal cycles were carried out as follows: 94 °C for 5 min, followed by 25 cycles of 94 °C for 1 min, 45 °C for 1 min, and 72 °C for 2 min. The PCR product was cloned into the pMD18-T vector (Takara, Japan) and sequenced using the chain-terminating method with an ABI automated sequencer (model no. 377). The cDNA from *Nicotiana tabacum* L. was termed NtCP56 (GenBank accession no. EU429306).

Phylogenetic analysis

A phylogenetic tree between NtCP56 and the other major classes of plant CPs was constructed with MEGA v.2.1 by using the Neighbor-Joining (NJ) method with bootstrapping 500 replicates.

Northern blotting analysis

For the RNA gel blot analysis, 25 μ g of RNA was fractionated by gel electrophoresis in a 1.2% formaldehyde agarose gel, then transferred to a Hybond-N⁺ membrane (Amersham, UK) after UV cross-linking (Yu *et al.*, 2003). The gene-specific DNA probe was prepared by amplifying a 356 bp 5' cDNA fragment of NtCP56. The probe was labelled using a DIG-primer labelling kit according to the instructions of the manufacturer (Roche, Germany). Membrane-bound DNA was prehybridized for 1 h and the hybridization of probes took place overnight at 42 °C. Two subsequent stringency washes were performed at 25 °C for 10 min and at 65 °C for 30 min. The first washing solution contained 0.1% SDS (w/v), 2 \times SSC, and the second 0.1% SDS (w/v), 0.5 \times SSC. Kodak X-Omat AR-5 films were exposed to the blots (Amersham, UK).

Expression, purification, renaturation, and sequencing of recombinant NtCP56

The open reading frame minus the first 60 bp of the cDNA encoding NtCP56 was amplified. The PCR product was digested with *Kpn*I and *Sac*I and subcloned into the plasmid expression vector pET30a (Novagen Inc., Madison, WI, USA). Fresh colonies of BL21 (DE3) bacteria transformed with plasmid pET30a/NtCP56 were incubated overnight at 37 °C. Overnight cultures were diluted to 1:100 in a 1.0 l fermenter and grown to mid-log phase. 0.4 mM IPTG was added and cultures were incubated for 3–5 h following induction. Bacterial pellets were harvested and stored at –70 °C (Smith and Gottesman, 1989).

The resuspended bacteria were disrupted by cold sonication in a Heat Systems-Ultrasonic cell disruptor at 200 W (120 \times 6 s) after suspension in the buffer; the insoluble protein fractions were isolated by centrifugation for 10 min at 12 000 *g*. The pellets were washed twice in the same buffer. The insoluble protein pellets were resuspended

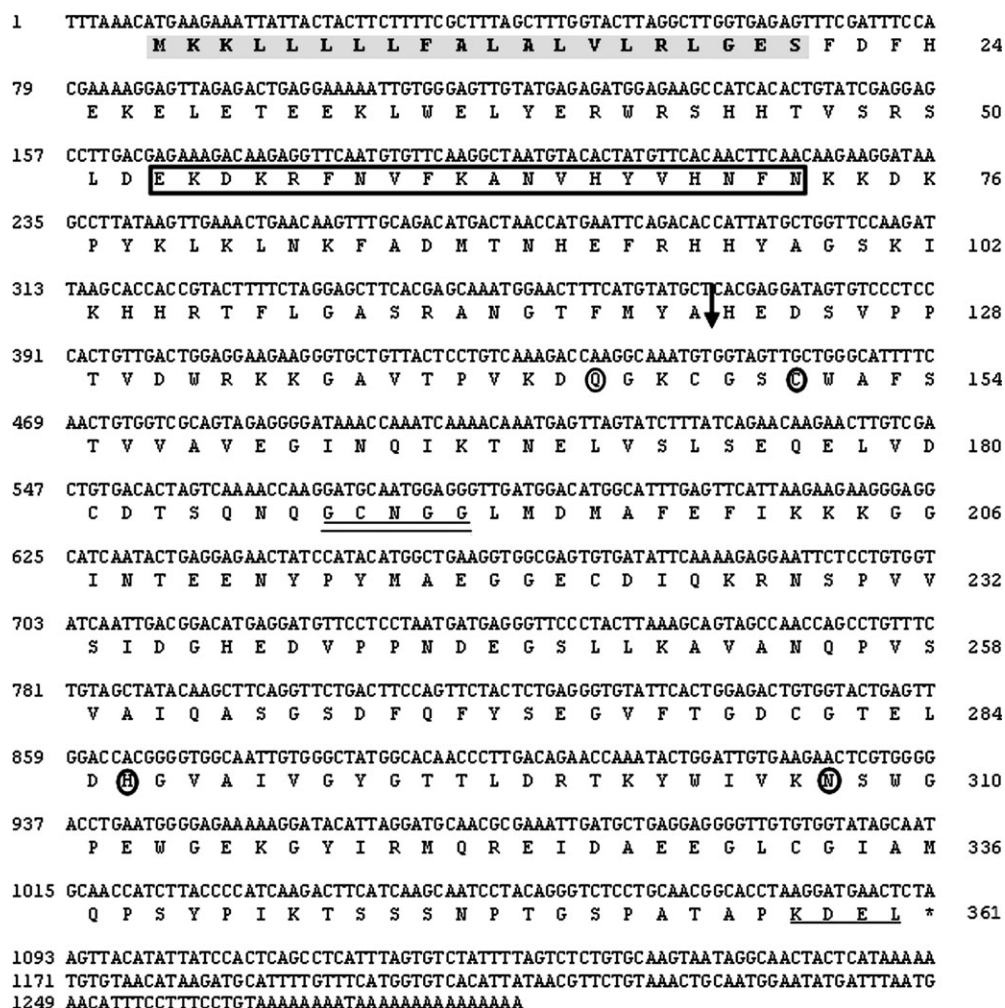


Fig. 1. Nucleotide and deduced amino acid sequence of NtCP56. The signal peptide (M1-S20) of NtCP56 is shown in bold. The catalytic triad Cys, His, and Asn and also the Glu active site residue are circled. The GCNNG motif is double-underlined. ERFNIN motif (E53-N72) of NtCP56 is shown in a rectangular box and the KDEL (K358-L361) motif is single underlined. The arrow indicates the site of auto-hydrolysis. Numbers on the left and right margins represent nucleotide and deduced amino acid sequences, respectively.

using 3 ml g⁻¹ bacteria of buffer A (10 mM TRIS-HCl, 8 M urea, 50 mM NaH₂PO₄ and the pH adjusted to 8.0 using NaOH) and incubated at 28 °C for 1 h.

The resuspended solution were added to the Ni²⁺-NTA resin (4 ml) which was equilibrated by buffer A in a 15 ml conical tube and gently shaken at 4 °C for 1 h. Then the bottom cap was removed and the flow-through was collected. The column was washed with 20 column volumes of washing buffer B (10 mM TRIS-HCl, 8 M urea, 50 mM NaH₂PO₄ and the pH adjusted to 6.3 using HCl) and buffer C (10 mM TRIS-HCl, 8 M urea, 50 mM NaH₂PO₄ with the pH adjusted to 5.9 using HCl). The over-expressed protein was then eluted in 4 ml fractions in buffer D (10 mM TRIS-HCl, 8 M urea, 50 mM NaH₂PO₄ and the pH adjusted to 4.5 using HCl) (Handbook for high-level expression and purification of 6× His-tagged proteins, Novagen).

Denatured proteins in a dialysis bag were immersed into a 50-fold volume renaturation buffer (50 mM potassium phosphate, pH 10.7, 5 mM EDTA, 1 mM reduced glutathione, 0.1 mM oxidized glutathione, 6 M/4 M/2 M/0

M urea) (Smith and Gottesman, 1989) and stirred at 4 °C for 3–4 h. Then the buffer was refreshed in turn by a reduction of urea.

Recombinant NtCP56 was processed to a mature form at pH 3.0 with cysteine (Troen *et al.*, 1988). NH₂-terminal amino acid sequencing was accomplished by the 2,4-dinitro-1-fluorobenzene (DNFB) method.

Enzyme assays and kinetic studies

Enzyme activity of purified recombinant NtCP56 was routinely assayed with casein (Sigma, USA) as substrate, using the method described by Sasaki *et al.* (1977) with minor modifications. Assays with up to 1% casein were carried out at 37 °C and pH 7.2, following 30 s activation incubation at 37 °C, pH 3.0, as described by previous reports (Troen *et al.*, 1988). Casein (1 ml of 0.04–1% solution in a phosphate-buffered saline containing 6 mM cysteine) was mixed with recombinant processed protein and the tubes were incubated at 37 °C for 3 min. After

incubation, the proteins were precipitated by the addition of 2 ml of a 20% tricarboxylic acid buffer and centrifuged at 12 000 *g* for 10 min. The absorbance of the clear supernatant was measured by a spectrophotometer at 275 nm. Controls, such as substrate and enzyme blanks, were included to obtain a background absorbance that was subtracted from the enzyme-containing tubes to yield a specific degradation of casein. The following two buffers were used: 100 mM sodium acetate (pH 3.0–5.0) and 100 mM sodium phosphate (pH 6.0–8.0). Kinetic constants were analysed by Matlab software. Leupeptin (100 μ M), antipain (100 μ M), and tosyl-L-lysine chloromethyl ketone (TLCK) (100 μ M) were used as inhibitors.

Anti-sense plasmid construction and generation of transgenic plants

The 356 bp fragment of NtCP56 near the 5' UTR was amplified and a *Bam*HI and *Xba*I restriction enzyme site was created by PCR. After digestion with *Bam*HI and *Xba*I restriction enzymes, the fragment was cloned into pBI121 (Clontech, USA) in an anti-sense orientation to create the 35S-anti-NtCP56 fusion construct. The fusion construct plasmid was first introduced into *Agrobacterium tumefaciens* LBA4404 and verified by using PCR. The empty vector (pBI121) was used as a control. Tobacco plant transformation from leaf explants was carried out as described by Lu *et al.* (2003).

Histological study

Anthers of control and transgenic tobacco were collected in a series of development stages. Anthers were fixed with FAA (5% formalin, 5% acetic acid, and 70% ethanol, by vol.) for 24 h under negative pressure and dehydrated in a set of increasing solutions (85%, 95%, and 100% by vol.), then decolourized in solutions of 50% dimethylbenzene and 50% ethanol (v/v) and dimethylbenzene and finally embedded in paraffin. Paraffin-embedded anthers were cut to a thickness of 10 μ m and stained with safranin and a fast green solution.

Results

Isolation and characterization of NtCP56

A 1289 bp cDNA was cloned from *Nicotiana tabacum* L. anthers. The cDNA contained a 1083 bp open reading frame (ORF), 7 bp of 5' untranslated region (UTR), and a 199 bp 3' UTR. The ORF encodes a peptide of 361 residues which has a predicted molecular mass of 43 kDa (Fig. 1). This cDNA shows 36–83% similarity with other KDEL-tailed plant cysteine proteinases, including the *Solanum lycopersicum* KDEL-tailed cysteine endopeptidase, *Ricinus communis* Cys-EP (Schmid *et al.*, 1998), *Phaseolus vulgaris* EP-C1 (Tanaka *et al.*, 1991), *Glycine max* CysP1 (Ling *et al.*, 2003), and *Vigna mungo* sulphhydryl endopeptidase SH-EP (Akasofu *et al.*, 1989). Then the cDNA is

named as NtCP56. A phylogenetic analysis was performed using the MEGA v.2.1 programme (Fig. 2). The NtCP56 protein is located in the group which contains KDEL-tailed family members. Given this analysis, the NtCP56 protein can be classified as belonging to the subfamily C1A of papain-like cysteine proteinases (MEROPS peptidase database, <http://merops.sanger.ac.uk>; Rawlings *et al.*, 2008).

The expression of the NtCP56 gene in different tissues of tobacco is determined using Northern blotting (Fig. 3).

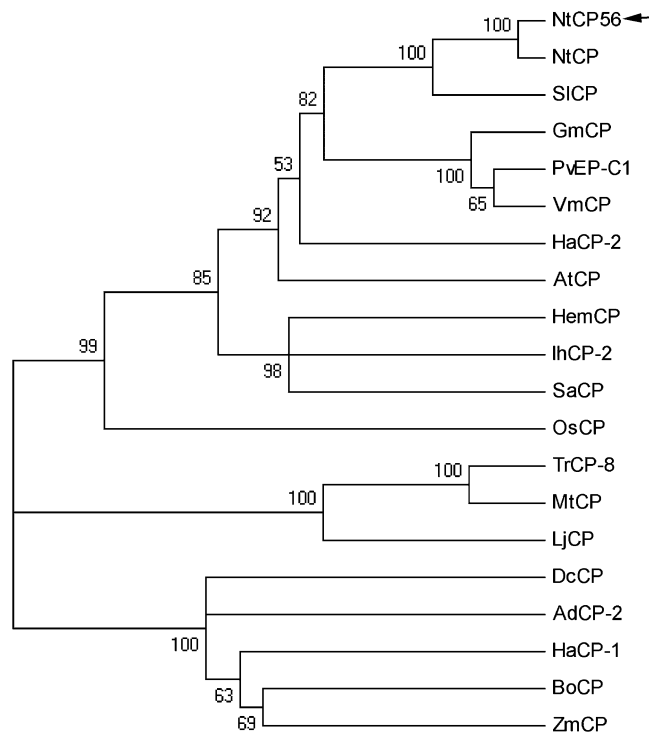


Fig. 2. Phylogenetic tree showing relationships between NtCP56 and other plant CPs. Bootstrap values from a sample of 500 replicates are shown on each branch using the following sequences NtCP (*Nicotiana tabacum*, GenBank accession no. AAW78660); SICP (*Solanum lycopersicum*, GenBank accession no. ABV22590); GmCP (*Glycine max*, GenBank accession no. BAC77523); PvEP-C1 (*Phaseolus vulgaris*, GenBank accession no. CAA40073); VmCP (*Vigna mungo*, GenBank accession no. P12412); HaCP (*Helianthus annuus*, GenBank accession no. BAC75924); AtCP (*Arabidopsis thaliana*, GenBank accession no. NP_568722); HemCP (*Hemerocallis* hybrid cultivar, GenBank accession no. P43156); lhCP (*Iris x hollandica*, GenBank accession no. AAR92155); SaCP (*Sandersonia aurantiaca*, GenBank accession no. AAD28477); OsCP (*Oryza sativa*, GenBank accession no. AAD20453); TrCP (*Trifolium repens*, GenBank accession no. AAP32196); MtCP (*Medicago truncatula*, GenBank accession no. AAQ63885); LjCP (*Lotus japonicus*, GenBank accession no. BAF56430); DcCP (*Daucus carota*, GenBank accession no. BAD29956); AdCP (*Actinidia deliciosa*, GenBank accession no. ABQ10200); HaCP (*Helianthus annuus*, GenBank accession no. BAC75923); BoCP (*Brassica oleracea*, GenBank accession no. AAL60580); ZmCP (*Zea mays*, GenBank accession no. NP_001104879).

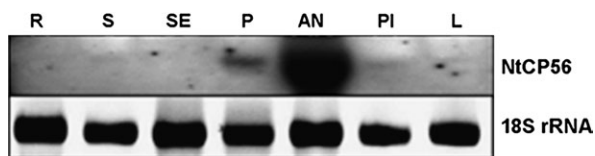


Fig. 3. Northern hybridization showing NtCP56 expression in the tissue of *Nicotiana tabacum* L. NtCP56 mRNA abundance was analysed by using the 5' end of NtCP56 cDNA as a probe. Each lane was loaded with 20 μ g total RNA. R, root; S, stem; SE, sepal; P, petal; AN, anther; PI, pistil; L, leaf. To ensure equal sample abundance on gels 18S rRNA was used to monitor loading equivalence.

The result showed that the NtCP56 gene is strongly expressed in anthers and weakly expressed in petal, pistil, stem, and leaf.

Expression, purification, and sequencing of recombinant NtCP56 protein

The plasmid which encodes the full-length NtCP56 minus the signal peptide was efficiently expressed in *E. coli*. Three hours after induction with IPTG, recombinant NtCP56 represented approximately 20% of the total bacterial protein. The molecular weight of the purified recombinant NtCP56 was estimated to be approximately 43 kDa according to the SDS-PAGE analysis (Fig. 4, Lane 3). This 43 kDa protein did not have any determinable activity (data not shown). After renaturation, the purified protein could be processed into a 33 kDa mature form at pH 3.0 (Fig. 4, Lane 4). The NH₂-terminal amino acid sequence of processed recombinant NtCP56 was HEDSVPP which, with four additional NH₂-terminal amino acids (HEDS), was compared with the predicted autocatalysis.

Biochemical characterization of processed recombinant NtCP56

The steady-state kinetics of the enzyme was studied in assays with various concentrations of casein as substrates. Processed recombinant NtCP56 obeyed the Michaelis-Menten-type of kinetics towards casein. The K_m (2.20 ± 0.04 mg ml⁻¹) and V_{max} (11.07 ± 0.23 μ g·ml⁻¹ min⁻¹) values for the recombinant enzyme are comparable to those reported for papain-like proteins from different species (Table I). The activity of the recombinant NtCP56 is pH-dependent. The highest activity was observed at pH 6.5, close to the optimal pH 7.0 of commercial papain (Hoover and Kokes, 1946). The enzyme still showed 53% of its maximum activity between pH 3.5 to pH 7.5, but only 28% at pH 8.5 (Fig. 5A), which was similar to the results of a previous report (Smith and Gottesman, 1989). The optimum temperature for NtCP56 activity was at 30 °C. The processed recombinant NtCP56 retained more than 50% of its maximum enzymatic activity between 20–60 °C (Fig. 5B).

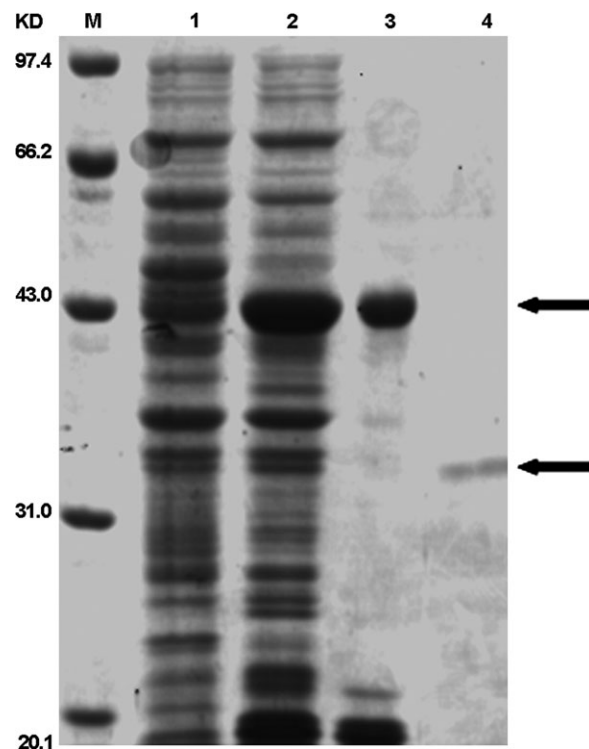


Fig. 4. Purification of recombinant NtCP56 and processed protein. Protein samples were subjected to SDS-polyacrylamide gel electrophoresis under reducing conditions and stained with Coomassie Brilliant Blue R-250. Lane M, molecular mass markers with the sizes shown on the left in kDa. Lane 1, induced BL21 (DE3)+PET30a whole cell lysate. Lane 2, induced BL21 (DE3)+pET30a/NtCP56 whole cell lysate. Lane 3, 43 kDa purified recombinant NtCP56. Lane 4, 33 kDa processed mature recombinant NtCP56.

Table 1. Enzymatic properties of recombinant NtCP56

	Recombinant NtCP56	Papain ^a	Fruit bromelain ^b
K_m (mg ml ⁻¹)	2.20 ± 0.04	1.4709	0.7
V_{max} (μ g ml ⁻¹ min ⁻¹)	11.07 ± 0.23	5.5252	134.2
K_{cat} (min ⁻¹)	1.025		
K_{cat}/K_m (ml mg ⁻¹ min ⁻¹)	0.46		
Inhibitor action (%)			
None			
Leupetin (100 μ M)	45.76%		
Antipain (100 μ M)	18.60%		
Tosyl-L-lysine chloromethyl ketone (TLCK) (100 μ M)	11.02%		

^a Xiao (2005).

^b Zhao *et al.* (1999).

Phenotypic analyses of anti-NtCP56 transgenic tobacco

In order to discover the function of NtCP56 in anther development, anti-NtCP56 transgenic tobaccos were obtained. In order to ascertain whether or not the *NtCP56*

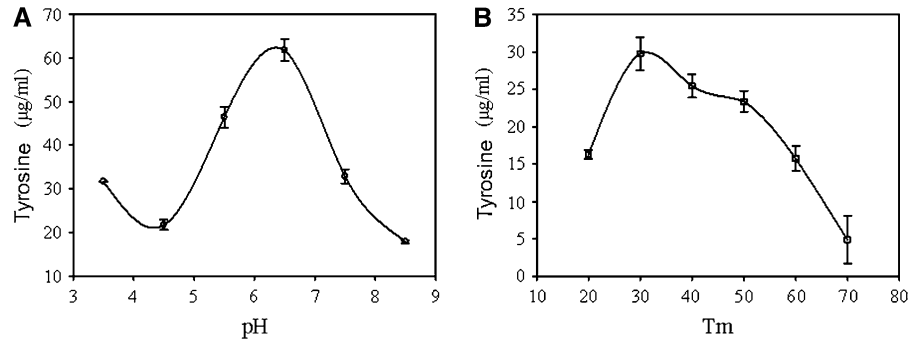


Fig. 5. Effect of pH and Tm on the enzyme activity. (A) pH stability of processed recombinant NtCP56. (B) Thermal stability of processed recombinant NtCP56.

gene plays a crucial role in the growth and development of tobacco plants, the root length and plant height of transgenic and control tobacco plants were measured each week. All 14 transgenic plants are phenotypically indistinguishable from the control plants under normal growing conditions. Compared with the control, the mRNA level of NtCP56 was reduced by 45–67% (Fig. 6A, B).

Anther development was concisely divided into six stages based on a previous classification of anther development (Feng *et al.*, 2001). During the microspore mother cell (MMC) stage, there was no detectable morphological difference in anthers between the control and anti-sense transgenic plants. Normal microsporocytes, an epidermis, an endothecium, a middle layer, and a tapetum were found both in controls and in transformation anthers (Fig. 7A, B). Up to the tetrad stage, there was still no obvious difference in anther cellular morphology between the control and anti-sense transformations. During this period, both in the control and in transgenic anthers, microsporocytes underwent meiosis to form tetrads of four haploid microspores. The tapetal cells had begun to differentiate and their cytoplasm became deeply stained (Fig. 7C, D) (Li *et al.*, 2006). Subsequently, the transgenic anthers had detectable morphological abnormalities. At the young microspore stage, in control anthers, microspores were released from tetrads and dispersed independently in locules (Fig. 7E). Partial tapetal cells were degenerated and the tapetum became thin (Fig. 7M). By contrast, in transformation anthers the tapetum remained relatively thick (Fig. 7N). Next, from the young pollen stage to the pollen mitosis stage, in control anthers, the uninucleate pollen developed to dianucleate pollen through mitotic divisions and dispersed in locules (Fig. 7G, I). Meanwhile, almost all tapetal cells had differentiated and degenerated (Fig. 7O). However, in the anti-NtCP56 transgenic anthers, abnormal microspores collapsed and remained undeveloped following the release from tetrads (Fig. 7H, J). Partial tapetum was still retained during this period (Fig. 7P). At the mature pollen stage, control pollen grains were full of starch, lipids, and nutrients (Fig. 7K) and the tapetum was fully degenerated (Fig. 7Q). By contrast, in transformation anthers, many pollen grains had become abnormal and did not mature (Fig. 7L, R). Large numbers of stagnated,

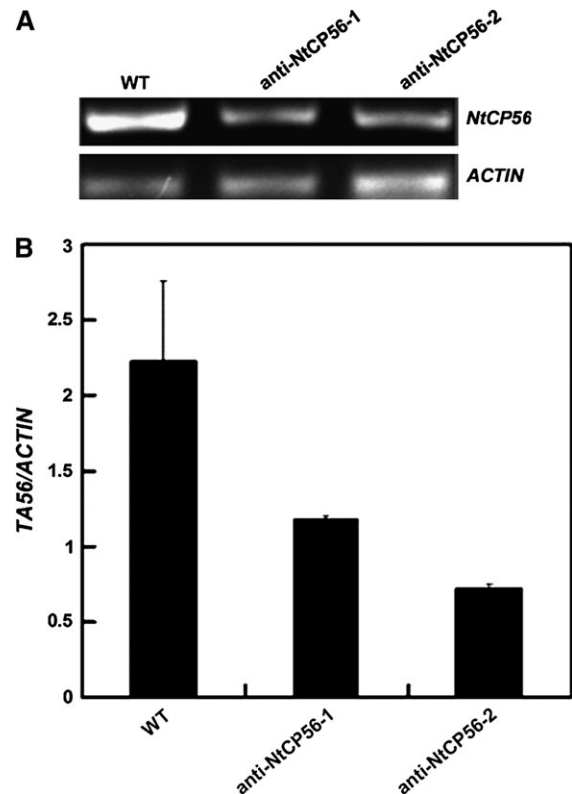


Fig. 6. Relative expression level of NtCP56 mRNA in control, anti-NtCP56-1 and anti-NtCP56-2 tobaccos. (A) Results of the semi quantity RT-PCR by 1% agarose. (B) Quantification of the result of the semi quantity RT-PCR.

deformed, and aggregated pollen grains were observed, and particularly in the corners of the locules (Fig. 7T). Most of the pollen grains did not stain densely, which suggested immaturity. In addition, from the young pollen stage, the cells of the endothecium gradually became fibrotic in control anthers. By contrast, in the transformation of anti-NtCP56, the endothecium cells remained loosely arrayed and did not have clear fibrotic cell walls.

In order to validate the sterility of transformation pollen grains further, the germination rate of pollen grains *in vitro* were determined. The germination rate of pollen grains is 41.3% of transgenic lines (102 in 247) and 5.3% of control

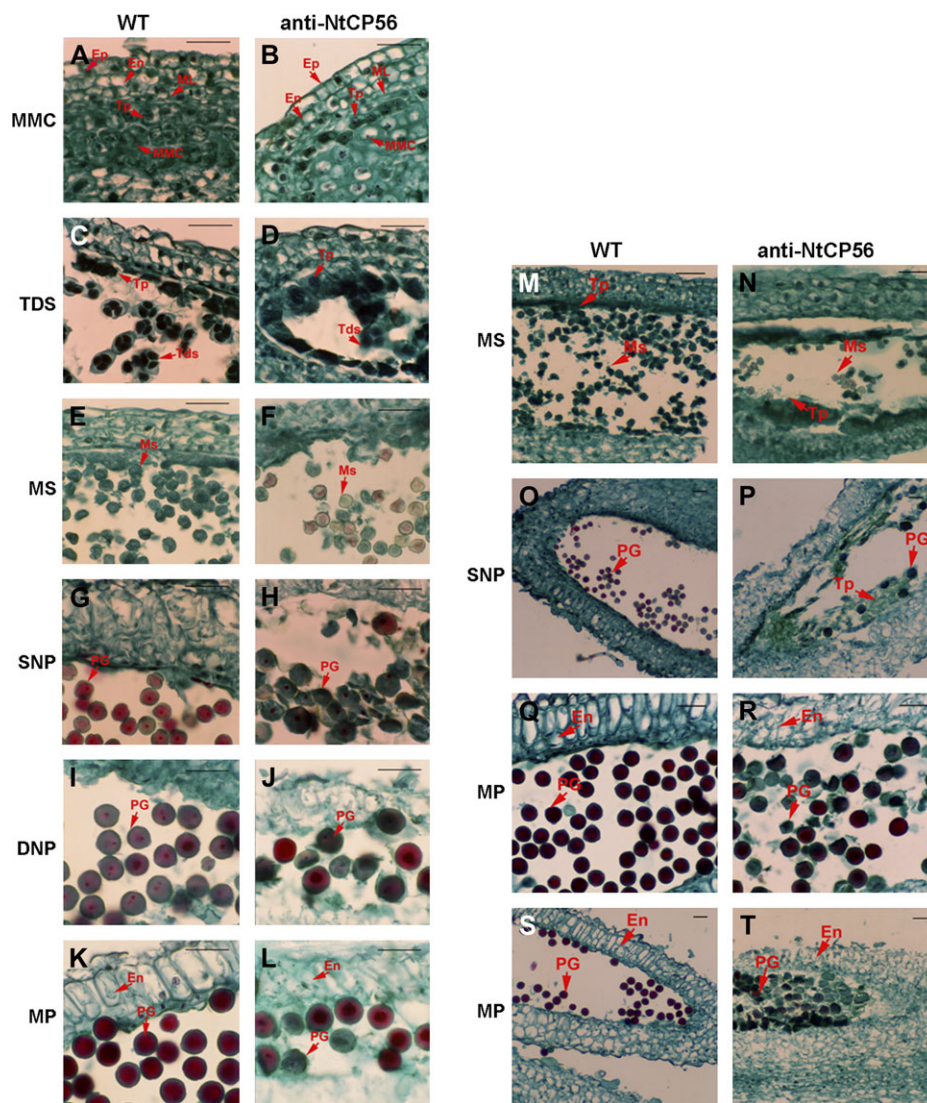


Fig. 7. Lengthways section comparison of the anther development of the control and the anti-NtCP56 transgenic tobaccos. Six stages of anther development in the control and the corresponding stages of development in the anti-NtCP56 transformation were compared. Control sections are shown in (A), (C), (E), (G), (I), (K), (M), (O), (Q), (S), and (B), (D), (F), (H), (J), (L), (N), (P), (R), and (T) show corresponding anti-NtCP56 sections. (S) and (T) show the locules corner sections of control and anti-NtCP56 transgenic tobaccos in the last stage. Tds, tetrads; PG, pollen grains; Ms, microspore; Tp, tapetum; Ep, epidermis; En, endothecium.; ML, middle layer. (A, B) The MMC stage; (C, D) The TDS stage; (E, F, M, and N) The young microspore stage; (G, H, O, P) The single nuclear pollen stage; (I, J) The dinuclear pollen stage; (K, L, Q, R, S, T) The mature pollen stage. (Bars=50 μ m).

lines (17 in 318). The germination rate of control lines is about eight times higher than that of transgenic lines. This result showed most transgenic pollen grains were abnormal in development.

These results suggested that the suppression of NtCP56 resulted in defects in pollen grain development.

Discussion

In this study, a gene from the tobacco anther, which is coded for a cysteine protease named NtCP56 has been isolated and characterized. Given the alignment to other plant papain-like cysteine proteases, the sequence of

NtCP56 was identified as a signal peptide that would be processed prior to enzyme activation. Like other papain-like cysteine proteases, three highly conserved catalytic residues Cys₁₅₀, His₂₈₆, and Asn₃₀₇ constitute the catalytic triad of cysteine proteases (Guerrero *et al.*, 1998). Another conserved residue, Gln₁₄₄, is involved in maintaining an active enzyme conformation (Kamphuis *et al.*, 1985). NtCP56 protein also contains a conserved non-contiguous ERFNIN motif (EX₃RX₃FX₂NX₃I/VX₃N) which is typical for cysteine proteases in the cathepsin L and H-like proteinases but not in cathepsin B-like proteinases (Wiederanders, 2003). A GCNGG motif was identified in NtCP56. With the exception of the central Asn (N) residue, this GCNGG motif is invariant in all ERFNIN proteinases

and also in the cathepsin B-like proteinases (Beyene *et al.*, 2006). In addition, a C-terminal tetrapeptide KDEL is present. It has been suggested that this is an endoplasmic reticulum (ER) retention signal (Fig. 1) (Guerrero *et al.*, 1998).

Foreign proteins overexpressed in *E. coli* frequently formed insoluble aggregates or inclusion bodies in the bacterial cytoplasm (Marston, 1986). When plasmid pET30a-NtCP56 was expressed in *E. coli*, the recombinant proteins quickly formed inclusion bodies and could only be solubilized in strong detergents such as 8 M urea. Denatured soluble protein was eluted at pH 4.5 after purification by Ni-NTA resin. Renaturation was carried out at neutral or alkaline pH to ensure recovery of the pro-enzyme in order not to be autocatalysed (Gal and Gottesman, 1986). This might be a result of the difference between the maturation conditions *in vitro* and in the plant (Vernet, 1990).

The NtCP56 gene is strongly expressed in anthers. This result is analogous to *OsCPI* reported in 2004, where its promoter was highly active in the anther, but not in other flower organs and vegetative organs of rice (Lee, 2004). This is different with most homologous CPs which have been reported and participated in processing and degradation of seed storage proteins (Shimada *et al.*, 1994; Toyooka *et al.*, 2000), fruit ripening (Alonso and Granell, 1995) and also in legume nodule development (Naito *et al.*, 2000). But it was similar to *OsCPI* isolated from rice, which is involved in pollen development (Lee *et al.*, 2004). It has been suggested that cysteine proteases have separate roles in different tissues (Lee *et al.*, 2004). RD21 (Yamada *et al.*, 2001), a cysteine protease that belongs to the papain family in *Arabidopsis*, has a role in the degradation of cellular proteins during leaf senescence, rather than in the degradation of seed storage protein after seed germination. Thus, it is speculated that NtCP56 is a novel cysteine protease which might be involved in anther development.

It has been demonstrated that the protease activity increased during anther development (DeGuzman and Riggs, 2000). This enhanced activity can be correlated with morphological and biochemical events during late microsporogenesis which require proteolytic enzymes. Pollen development is a dynamic process that involves protein synthesis required for the development (Lee *et al.*, 2004). In addition, the formation and degradation process of the tapetum which supplies nutritional sources to developing pollen was also vital to pollen development (Ku *et al.*, 2003).

An anti-sense approach was used to analyse the role of NtCP56 in pollen development. Until the tetrad stage, the anther development of transgenic plants was normal. However, after that, development of both tapetum and microspores was disrupted in anti-NtCP56 anthers. It suggested that NtCP56, as a cysteine protease, may participate in new protein synthesis for pollen development or aid in the degenerative process of the tapetum. It is either activating hydrolytic enzymes that degrade cellular macromolecules, or by functioning as a hydrolytic enzyme, or both (Koltunow *et al.*, 1990). At present, the mechanism of NtCP56 during pollen development remains to be explained.

Acknowledgements

Financial source: National Natural Science Foundation of China (Grant No.30070612). This work was supported by grant from the National Basic Research Program of China (2009CB119104). The authors also thank Jing-chuan Du, Xu Wang (The Chinese Academy of Agriculture Sciences), Nan Qiao and members of the Botany laboratory for support during data analysis.

References

- Akasofu H, Yamauchi D, Mitsuhashi W, Minamikawa T.** 1989. Nucleotide sequence of cDNA for sulfhydryl-endopeptidase (SHEP) from cotyledons of germinating *Vigna mungo* seeds. *Nucleic Acids Research* **17**, 6733.
- Albrecht C, Russinova E, Hecht V, Baaijens E, de Vries S.** 2005. The *Arabidopsis thaliana* SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis. *The Plant Cell* **17**, 3337–3349.
- Alonso JM, Granell A.** 1995. A putative vacuolar processing protease is regulated by ethylene and also during fruit ripening in citrus fruit. *Plant Physiology* **109**, 541–547.
- Beyene G, Foyer CH, Kunert KJ.** 2006. Two new cysteine proteinases with specific expression patterns in mature and senescent tobacco (*Nicotiana tabacum* L.) leaves. *Journal of Experimental Botany* **57**, 1431–1443.
- Colcombet J, Boisson-Dernier A, Ros-Palau R, Vera CE, Schroeder JI.** 2005. *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are essential for tapetum development and microspore maturation. *The Plant Cell* **17**, 3350–3361.
- Davis TA.** 1966. Floral structure and stamens in *Bombax ceiba* L. *Journal of Genetics* **59**, 294–299.
- DeGuzman R, Riggs CD.** 2000. A survey of proteinases active during meiotic development. *Planta* **210**, 921–924.
- Feng JH, Lu YG, Liu XD, Xu XB.** 2001. Pollen development and its stages in rice (*Oryza sativa* L.). *Chinese Journal of Rice Science* **15**, 21–28.
- Gal S, Gottesman MM.** 1986. The major excreted protein of transformed fibroblasts is an activable acid-protease. *Journal of Biological Chemistry* **261**, 1760–1765.
- Guerrero C, Calle M, Reid MS, Valpuesta V.** 1998. Analysis of the expression of two thioprotease genes from day lily (*Hemerocallis* spp.) during flower senescence. *Plant Molecular Biology* **36**, 565–571.
- Hoover SR, Kokes ELC.** 1946. Effect of pH upon proteolysis by papain. *Journal of Biological Chemistry* 199–207.
- Ito T, Shinozaki K.** 2002. The *MALE STERILITY1* gene of *Arabidopsis*, encoding a nuclear protein with a PHD-finger motif, is expressed in tapetal cells and is required for pollen maturation. *Plant and Cell Physiology* **43**, 1285–1292.
- Jung KH, Han MJ, Lee YS, Kim YW, Hwang I, Kim MJ, Kim YK, Nahm BH, An G.** 2005. Rice Undeveloped Tapetum1 is a major regulator of early tapetum development. *The Plant Cell* **17**, 2705–2722.
- Kamphuis IG, Drenth J, Baker EN.** 1985. Comparative studies based on the high-resolution structures of papain and actinidin, and

on amino acid sequence from cathepsins B and H, and stem bromelain. *Journal of Molecular Biology* **182**, 317–329.

Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RB. 1990. Different temporal and spatial gene expression patterns occur during anther development. *The Plant Cell* **2**, 1201–1224.

Ku S, Yoon H, Suh HS, Chung YY. 2003. Male-sterility of thermosensitive genic male-sterile rice is associated with premature programmed cell death of the tapetum. *Planta* **217**, 559–565.

Lee S, Jung KH, An G, Chung YY. 2004. Isolation and characterization of a rice cysteine protease gene, OsCP1, using T-DNA gene-trap system. *Plant Molecular Biology* **54**, 755–765.

Li N, Zhang DS, Liu HS, et al. 2006. The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *The Plant Cell* **18**, 2999–3014.

Li QZ, Li XG, Bai SN, Lu WL, Zhang XS. 2002. Isolation of HAG1 and its regulation by plant hormones during *in vitro* floral organogenesis in *Hyacinthus orientalis* L. *Planta* **215**, 533–540.

Ling J, Kojima T, Shiraiwa M, Takahara H. 2003. Cloning of two cysteine proteinases genes: *CysP1* and *CysP2*, from soybean cotyledons by cDNA representational difference analysis. *Biochimica et Biophysica Acta* **1627**, 129–139.

Lu H, Zeng QY, Zhao YL, Jiang XN. 2003. Xylem-specific expression of a GRP1.8 promoter::4CL gene construct in transgenic tobacco. *Plant Growth Regulation* **41**, 279–286.

Marston FA. 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochemical Journal* **240**, 1–12.

Mizuno S, Osakabe Y, Maruyama K, Ito T, Osakabe K, Sato T, Shinozaki K, Shinozaki KY. 2007. Receptor-like protein kinase 2 (RPK 2) is a novel factor controlling anther development in *Arabidopsis thaliana*. *The Plant Journal* **50**, 751–766.

Naito Y, Fujie M, Usami S, Murooka Y, Yamada T. 2000. The involvement of cysteine proteinase in the nodule development in Chinese milk vetch infected with *Mesorhizobium huakuii* subsp. *rengei*. *Plant Physiology* **124**, 1087–1096.

Nonomura K, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, Kurata N. 2007. A germ cell specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. *The Plant Cell* **19**, 2583–2594.

Rawlings ND, Tolle DP, Barrett AJ. 2004. MEROPS, the peptidase database. *Nucleic Acids Research* **32**, 160–164.

Sanders PM, Bui AQ, Weterings K, McIntire KN, Hsu YC, Lee PY, Truong MT, Beals TP, Goldberg RB. 1999. Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sexual Plant Reproduction* **11**, 297–322.

Sasaki M, Minakata K, Yamamoto H, Niwa M, Kato T, Ito N. 1977. A new serum component which specifically inhibits thiol proteinases. *Biochemical and Biophysical Research Communications* **76**, 917–924.

Schmid M, Simpson D, Kalousek F, Gietl C. 1998. A cysteine endopeptidase with a C-terminal KDEL motif isolated from castor bean endosperm is a marker enzyme for the ricinosome, a putative lytic compartment. *Planta* **206**, 466–475.

Scott RJ, Spielman M, Dickinson HG. 2004. Stamen structure and function. *The Plant Cell* **16**, 46–60.

Shimada T, Hiraiwa N, Nishimura M, Hara-Nishimura I. 1994. Vacuolar processing enzyme of soybean that converts proproteins to the corresponding mature forms. *Plant and Cell Physiology* **35**, 713–718.

Smith SM, Gottesman MM. 1989. Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli*. *Journal of Biological Chemistry* **264**, 20487–20495.

Solomon M, Belenghi B, Delledonne B, Menachem E, Levine A. 1999. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *The Plant Cell* **11**, 431–443.

Tanaka T, Yamauchi D, Minamikawa T. 1991. Nucleotide sequence for an endopeptidase (EP-C1) from pods of maturing *Phaseolus vulgaris* fruits. *Plant Molecular Biology* **16**, 1083–1084.

Toyooka K, Okamoto T, Minamikawa T. 2000. Mass transport of a proform of a KDEL-tailed cysteine proteinase (SH-EP) to protein storage vacuoles by endoplasmic reticulum-derived vesicle is involved in protein mobilization in germinating seeds. *Journal of Cell Biology* **148**, 453–563.

Troen BT, Ascherman D, Atlas D, Gottesman MM. 1988. Cloning and expression of the gene for the major excreted protein of transformed mouse fibroblasts. *Journal of Biological Chemistry* **263**, 254–261.

Varnier AL, Mazeyrat-Gourbeyre F, Sangwan RS, Clement C. 2005. Programmed cell death progressively models the development of anther sporophytic tissues from the tapetum and is triggered in pollen grains during maturation. *Journal of Structural Biology* **152**, 118–128.

Vernet T, Tessier DC, Richardson C, Laliberte F, Khouri HE, Bell AW, Storer AC, Thomas DY. 1990. Secretion of functional papain precursor from insect cells. *Journal of Biological Chemistry* **265**, 16661–16666.

Wiederanders B. 2003. Structure–function relationships in class CA1 cysteine peptidase propeptides. *Acta Biochimica Polonica* **50**, 691–713.

Wilson ZA, Morroll SM, Dawson J, Swarup R, Tighe PJ. 2001. The *Arabidopsis* MALE STERILITY1 (*MS1*) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. *The Plant Journal* **28**, 27–39.

Xiao GP. 2005. Ultrasonic extraction technology of papain and its enzymatic properties. *Journal of Fujian Agriculture and Forestry University* **34**, 318–323.

Xu FX, Chye ML. 1999. Expression of cysteine proteinase during developmental events associated with programmed cell death in brinjal. *The Plant Journal* **17**, 321–327.

Yamada K, Matsushima R, Nishimura M, Hara-Nishimura I. 2001. A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing *Arabidopsis* leaves. *Plant Physiology* **127**, 1626–1634.

Yu Y, Steinmetz A, Meyer D, Brown S, Shena WH. 2003. The tobacco A-type cyclin, *Nicta*; CYCA3;2, at the nexus of cell division and differentiation. *The Plant Cell* **15**, 2763–2777.

Zhao WL, Wu W, Li CM, Wang XJ. 1999. Kinetic study of fruit bromelain. *Journal of China Agricultural University* **4**, 11–13.