# Molecular cloning and primary sequence analysis of a gene encoding a putative chitinase gene in *Brassica oleracea vat. capitata*

TANG GUO QING, YONG YAN BAI, SHI WEI LOO Shanghai Institute of Plant Physiology, Chinese Academy of Sciences, 300 Fenglin Road Shanghai, China 200032.

## ABSTRACT

Chitinase, which catalyzes the hydrolysis of the  $\beta$ -1, 4-acetyl-D-glucosamine linkages of the fungal cell wall polymer chitin, is involved in inducible plants defense system. By construction of cabbage (Brassica oleracea var. *capitata*) genomic library and screening the library with pRCH8, a probe of rice chitinase gene fragment, a chitinase genomic sequence was isolated. The complete nucleotide sequence of the putative cabbage chitinase gene (cabch29) was determined, with its longest open reading frame (ORF)encoding a polypeptide of 413 aa. This polypeptide consists of a 21 aa N-terminal signal peptide, two chitin-binding domains different from those of other classes of plant chitinases, and a catalytic domain. Homology analysis illustrated that this cabch29 gene has 58.8% identity at the nucleotide level with the pRCH8 ORF probe and has 50% identity at the amino acid level with the catalytic domains of chitinase from bean, maize and sugar beet. Meanwhile, several kinds of cis- elements, such as TATA box, CAAT box, GATA motif, ASF-1 binding site, wound-response elements and AATAAA, have also been discovered in the flanking region of cabch29 gene.

**Key words:** Cloning, sequence, cabbage, chitinase gene.

Molecular cloning of chitinase gene

## INTRODUCTION

Plants have developed several biochemical defense mechanisms in response to pathogens and abiotic stress. Following pathogen attack, plant synthesize phenyl-propaniod products such as lignin, low mol. wt. antimicrobial compounds known as phytoalexins, and several defense-related proteins. Among these proteins are "pathogenesis-related proteins" including the fungal cell wall degrading enzymes chitinase and  $\beta$ -1, 3-glucanase[1].

Endochitinase from higher plants catalyze the hydrolysis of chitin, a  $\beta$ -1, 4-1inked homopolymer of N-acetyl-D-glucosamine. The level of chitinase activity increases dramatically after invasion by fungal[2, 3], bacterial[4], or viral pathogens[5]. Although chitin does not exit in plant cells, it is a major component of the many fungal cell wall[6]. Purified plant endochitinase have antifungal activity against some fungi *in vitro*[7] and can act synergistically with  $\beta$ -1, 3-glucanases purified from plants to inhibit fungal growth[8]. Furthermore, the presence of pathogenesis-related proteins is associated with some hypersensitive response[5], and induced resistance[9]. These observations suggest that chitinases are part of a general disease resistance mechanism.

We have initiated to isolate and characterize the chitinase gene (cabch29) in *Brassica oleracea* var.*capitata*, a popular vegetable cabbage. Sequence homology analysis indicated that the cabch29 probably should be assigned to a new class of plant chitinase.

## MATERIALS AND METHODS

#### Construction and screening of genomic library

The Brassica olerecea var. capitata genomic library was constructed according to standard procedure[10]. The library was screened by plaque hybridization using a random-primer labeling rice DNA encoding a basic chitinase (probe pRCH8, a gift from Dr. Qun ZHU, Salk Institute, USA). Hybridization was carried out at  $42^{\circ}$ C in 50% formanide,  $5 \times$  SSPE,  $5 \times$  Denhart's solution, 0.2%(w/v) SDS. Filters were washed sequentially at low stringency ( $2 \times$  SSPE, 0.2% SDS for 30 min at room temperature and  $37^{\circ}$ C).

### Southern blotting and hybridization

Restriction mapping of a positive clone (lambda CH29) was carried out. DNA fragments from agarose gel was transferred to Hybond-N+nylon membranes as recommended by the manufacturer (Amersham). The probe labeling, hybridization and wash membrane were performed under the same conditions described for plaque hybridization.

#### Subcloning and DNA sequencing

As shown in figure 1, a 2.7 kb Sall/HindIII fragment from lambda EMBL4 genomic clone lambda CH29 was subcloned into pBluescriptIIKS(+). The nested deletion sequencing templates were prepared according to previously described methods[11]. DNA sequence was determined by the chain termination method[12] using T7 DNA polymerase Kit (Shanghai Promega. Co.).

## RESULTS

## Isolation of genomic clones and sequencing

A genomic library of cabbage nuclear DNA was screened by plaque hybridization using rice chitinase clone pRCH8 as the probe. 5 positive phages were obtained by screening  $8.2 \times 10^6$  plaques. These were purified and partially characterized by Southern blot analysis. Among these positive phages, lambda CH29 was chosen for detailed study because the signal of hybridization shown that the positive phage had single intact chitinase gene (Fig 1).

Upon digestion with Sall/HindIII, lambda CH29 yield a 2.7 kb fragment which contained the complete chitinase coding sequence. Sequentially, these fragment was subcloned into pBluescriptIIKS(+) to get pcabch29 (Fig 1).



Fig 1. Schematic representation of the cloning, subcloning and the strategy of sequencing. The 6.8 kb fragment was released from lambda CH29 by digestion with SalI and inserted in plasmid vector PSal 6.8. was digested with HindIII, self-ligated and generated pcabch29, the nested deletion sequencing templates were prepared from pcach29 by ApaI/XbaI double digestion, modified with ExoIII and S1 nuclease. Abbreviation are: A, ApaI. B, BamHI. H, HindIII. K, KpnI. S, SalI. X, XbaI.

The transcriptional orientation of gene was deduced by hybridizing differentially to probes specific to the 5' or 3' regions respectively of rice chitinase gene pRCH8.

Two direction nested deletion templates from pcabch29 were prepared and sequenced. The nucleotide sequence of relevant regions of lambda CH29 is shown in Fig 2.

## Characterization of nucleotide sequence of cabbage chitinase gene cabch29

By PC gene program analysis, we have discovered some putative cis-elements of flanking region and open reading flame (ORF) in the cabbage chitinase gene cabch29.



**Fig 2.** The complete nucleotide and deduced amino acid sequence of cabbage (*Brassica oleracea* var. *capitata*) chitinase gene cabch29. The length of the sequence is 2768 bp, and the longest putative ORF is 1242 bp encoding for 413 aa by computer-aided suggestion. The first nucleotide of the ORF was pointed as +1; the deduced code region and translation and transcription stop codes are shown in the bold characters; and some cis-elements, e.g. CAT, TATA boxes and transcription stop codes underlined.

#### 3'-and 5'-flanking region

Several sequences in the 5'-flanking region could, by analogy with other eukaryotic genes, be involved in transcriptional regulation. The nucleotide sequence CACCAT-GAG is in fair agreement with the consensus sequence proposed for the start of plant gene translation[13], ATG position was pointed at +1. The putative TATA box, TATATAAA, is upstream of the putative translation start at positions -100 to -93. The sequence CCAAT at position -129 is similar to the CAAT box of animal genes sometimes found upstream of the TATA box[14]. Some of these cis-elements are shown in Fig 2. Beside, many other putative cis-elements exist in the 5' flanking region of cabch29, which include: 1) some wound-response elements, such as AGC box and TCA motif [15], carrot extensin gene wound-response elements [16, 17], and elicitor-inducible PAL footprint. 2) tissue-specific elements such as ASF-1 binding site and GATA motif [18]. 3)elements maintaining transcription such as G box [19]. It is worthwhile to distinguish or identify functional cis-element from all those putative ones listed by further experiment.

In the 3'-flanking region, there are two AATAAA sequences at positions 72 and 84 downstream of the TAA translation stop, which are likely to be involved in processing for polyadenylation[20].

#### The primary structure of cabbage chitinase

The computer-aided analysis indicated that cabch29 contains 2768 bp with a 1242 bp nucleotide ORF coding for 413 aa and the nucleotide sequence CACCATGAG (marked with +1 in Fig.l) is probably the translation initiation code region.

Cabch29 polypeptide contains a hydrophobic putative signal peptide of 21 aa at the N-terminus, which has the structure expected for a signal peptide[21]. Cabch29 gene also contains hevein(chitin-binding domain)which is different from other plant chitinases. Cabch29 gene product contains two putative hevein domains (namely, hevein domain A and hevein domain B). Hevein domain A is of 43 aa long with 10 cysteines, and hevein domain B 35 aa with 8 cysteines. Different from class V chitinase in *Urtica dioica*, the distance between the two hevein domains within cabch29 gene product is about 90 aa, whereas that in *Urtica dioica* is 7 aa. Following hevein domain B is catalytic domain seperated by glycine and proline-rich spacer region of 11 aa.

#### Homology of the putative cabbage chitinase to other chitinase

The primary structure of the putative cabch29 gene product, deduced from the nucleotide sequence, was compared to basic and acidic chitinase from other plants. The degree of overall homology between putative cabch29 gene product and other plant chitinases is low, but the result of domain comparision has shown similarity to some extent and even a much higher one in certain hevein domain (Hvd). As shown in Fig 3, a much higher as sequence identity (70%-80%) was found between Hvd

Molecular cloning of chitinase gene

A of cabbage chitinase cabch29 and that of class I chitinase of rice chitinase Hvd 1 and bean Hvd 2. On the other hand, a much lower level of aa sequence identity (47-54%) is found between the cabch29 hevein domains (Hvd A, Hvd B) and these of other plant chitinase and chitin-binding proteins, such as tobacco (Hvd 3) and potato (Hvd 4) chitinase, rubber hevein (Hvd 5), potato win1 (Hvd 6), and win2 (Hvd 7), WGA-B (Hvd 8) and rice lectin (Hvd 9).

| Hvd | 1 | -EQCGSQAGGALCPNCLCCSQYGWCGSTSAYCGSGCQSQCSG-   |
|-----|---|---|
| Hvd | А | -EGCGSQAGGPVCPNCLCCSQFGWCGSTSDSCGGGCQSQCSRL   |
| Hvd | 2 | -EQCGRQAGGALCPGGNCCSQFGWCGSTTDYCGPGCQSQCGG-   |
| Hvd | 5 | -EQCGRQAGGKLCPNNLCCSQWGWCGSTDEYCSPDHNCQSNCKD- |
| Hvd | 6 | -QQCGRQKGGALCSGNLCCSQFGWCGSTPEFCSPSQGCQSRCTG- |
| Hvd | 7 | -QQCGRQRGGALCGNNLCCSQFGWCGSTPEYCSPSQGCQSQCTG- |
| Hvd | 3 | -EQCGSQAGGARCASGLCCSKFGWCGNTNDYCGPGN-CQSQCPG- |
| Hvd | 4 | -QNCGSQGGGKACASGQCCSKFGWCGNTNDYCGSGN-CQSQCPG- |
| Hvd | 8 | -QRCGEQGSGMECPNNLCCSQYGYCGMGGDYCGKGCQNGACW-   |
| Hvd | 9 | -QTCGKQNDGMICPHNLCCSQFGYCGLGRDYCGTGCQSGACC-   |
| Hvd | В | SQNCGCAPNLCCSQFGYCGSTDAYCGTGCRSGPCRS          |

Fig 3. Alignment of hevein-domain in several plant chitinases and chitin- binding proteins. Hvd A and Hvd B from cabbage chitinase gene cabch29. Hvd 1 from rice, Hvd 2 from bean, Hvd 3 from tobacco and Hvd 4 from potato chitinases. Hvd 5 from rubber, Hvd 6 and Hvd 7 from potato winl, win2, Hvd 8 from WGA-B and Hvd 9 from rice lectin are chlitin binding proteins.

The catalytic domain is another key domain of chitinase and there is some similarity between various chitinases. The level of aa sequence identity between catalytic domain of cabch29 gene and those of other species (such as bean, tobacco, barley, nettle and tomato chitinases) were found to be low (27%-36%). In contrast, a slightly higher aa sequence identity (54%-58%) was found between catalytic domain of cabch29 gene and those of class IV chitinases such as maize, bean and basic sugar-beet chitinases(data not shown).

## DISCUSSION

Up till now, at least five distinct classes of chitinases have been assigned in plants. By comparison of various chitinase catalytic domains, pairwise similarities between individual chitinase are apparent, and these have play an important role for the isolation of various chitinase from plant genomic libraries. The present data described the isolation and characterization of a cabbage gene, cabch29, putatively encoding a chitinase. The pcabch29 genomic DNA fragment is 2768 bp in length and containing 1284 bp of 5' flanking sequence. The longest ORF in cabch29 is of 1242 bp and coding for 413 aa peptide. The first 21 aa residues at the N-terminus show the characteristics of an eukaryotic signal peptide with a highly hydrophobic core and the predicted aa sequence near the cleavage site[21]. Like other basic chitinases from dicotyledons[22-25] and a monocotyledon, rice[26], the cabch29 gene product is synthesized with a signal peptide for the transport of polypeptide across the membrane into the endoplasmic reticulum.

However, there is a striking feature, in contrast to what have been reported before, that the longest putative ORF of cabch29 has two hevein domains (hevein domain A and B). The ca.43 aa proximal to the signal peptide encode a hevein domain A linked by a spacer region(ca. 90 aa) to the hevein domain B (35 aa). The function of this spacer region is not clear, maybe, it contains an intervening sequence since it is AT-rich (60%). All these need to be further verified by experiment such as S1 mapping et al.

The hevein domains of cabch29 chitinase is cysteine-rich, sharing sequence identity not only with chitinase, but also with other chitin-binding proteins, such as rubber hevein, wheat germ agglutin in isolectin, rice lectin and the products of the potato winl and win2 genes. Furthermore, some chitin-binding proteins contain multiple hevein domains, eg. wheat germ agglutin in isolectin. The latter can bind chitin and other polysaccharides with N-acetylated amino sugar groups, and contains four hevein domains (each with 41-43 aa) which are similarly folded with disulfide bonds in homologous positions, and each one can putatively form two finger-like structures for binding sugar residues. Therefore, it would be interesting to see whether the two putative hevein domains of cabch29 chitinase will have similar function.

Enzyme activity of chitinase is mediated by its catalytic domain. Comparing the portion of catalytic domain of cabch29 chitinase with that of other chitinases, the former has 53.6%, 54.8% and 58.5% identity at aa level with sugar-beet basic chitinase and class IV chitinases from maize and bean respectively, although cabch29 chitinases showed much lower aa sequence identity to other chitinases from classes I, II, III and V.

A c-terminal extension is present in all published class I chitinases from dicotyledons [24-26]. With a tobacco chitinase gene, it has been found that some c-terminal sequences are necessary for targeting the protein to vacuoles[27]. No c-terminal extension was found in the cabbage cabch29 chitinase. Plant chitinases are a diverse group of enzymes with respect to structure, cellular localization and enzymatic properties, but cabch29 is, structurally, unique for having a double hevein domain and the lack of c-terminal extension. Thus, it can be inferred that the cabbage cabch29 gene may putatively represent a new class of plant chitinase gene.

### Molecular cloning of chitinase gene

## REFERENCES

- [1] Loon V. Stress proteins in infected plants. In Kosuge T, Nester EW, eds, Plant-Microbe Interactions. 1989; Vol. 3 McGraw-Hill, New York, pp.198-237.
- [2] Kombrink E, Schroder M, Hahlbrock K. Several "pathogenesis related" proteins in potato are 1, 3-β-glucanases and chitinase. Proc Natl Acad Sci USA 1988; 85:782-6.
- [3] Mauch F, Hadwiger LE, Boller T. Ethylene: Symptom, not signal for the induction of chitinase and  $\beta$ -1, 3-glucanase in pea pods by pathogens and elicitors. Plant physiology 1984; **76**:607-11.
- Broekaert WF, Peumans WJ. Pectic polysaccharides elicit chitinase accumlation in tobacco. Physiol Plant 1988; 74:740-4.
- [5] Legrand M, Kauffmann S, Geoffroy P. Fritig B. Biological function of pathogenesis-related proteins: Four tobacco pathogenesis-related proteins are chitinase. Proc Natl Acad Sci USA 1987; 88:6750-4.
- [6] Bartnicki-Garcia S. Cell wall chemistry, Morphogenesis and taxonomy of fungi. Annu Rev Microbiol 1968; 22:87-108.
- [7] Roberys WK, Selitrennikoff CP. Plant and bacterial chitinase differ in antifungal activity. J Gen Microbiol 1988; 134:169-76.
- [8] Mauch F, Mauch-Mani B, Boller T. Antifungal hydrolases in pea tissue II. Inhibition of fungal growth by combinations of chitinase and  $\beta$ -1, 3-glucanase. Plant Physiol 1988; **88**:936-42.
- [9] Tuzun S, Nageswara R, Vgeli U, Schardl C, Kuc J. Induced systemic resistance to blue mold: Early induction and accumulation of  $\beta$ -1, 3-glucanase, chitinase and other pathogenesis-related proteins (b-proteins) in immunized tobacco. Phytopathology 1989; **79**:979-83.
- [10] Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning, a laboratory manual, second edition, 1989; Cold Spring Harbr Laboratroy Press.
- [11] Henikoffs. Unidirectional digestion with exonuclease III creates targeted break points for DNA sequencing. Gene 1984; 28:351-9.
- [12] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 1977; 74:5463-8.
- [13] Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA. Selection of AUG initiation codons differs in plants and animals. EMBO J 1987; 6:43-8.
- [14] Breathnach R, Chambon P. Organization and expression of eukaryotic split genes coding for protein. Annu Rev Biochem 1986; 50:349-83.
- [15] Hart CM, Nagy F, Meins F Jr. A 61 bp enhancer element of tobacco  $\beta$ -1, 3-glucanase B gene intereacts with one or more regulated nuclear proteins. Plant Mol Biol 1993; **21**:121-9.
- [16] Lois R, Dietrich A, Hahlbrock K, Schulz W. A phenylalamine ammonia-lyase gene from parsley: structure, regulation and indentification of elicitor and light response cis-acting elements. EMBO J 1984; 8:1641-8.
- [17] Holdsworth MJ, Laties GG. Site-specific binding of a nuclear factor to the carrot extensin gene is influenced by both ethylene and wounding. Planta 1989; 180:74-81
- [18] Lam E, Chua NH. ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and conserved GATA motif in cab promoter. The Plant Cell 1989; 1:1147-56.
- [19] Datta N, Cashmore AR. Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. The Plant Cell 1989; 1:1069-77
- [20] Joshi CP. Putative polyadenylation signals in nuclear genes for higher plants: compilation and analysis. Nucl Acids Res 1987; **15**:9627-40.
- [21] Von Heijne G. Patterns of amino acids near signal sequence cleavage sites. Eur J Biochem 1983; 133:17-21.
- [22] Broglie KE, Gaynor J J, Broglie RM. Ethylene-regulated gene expression: molecular cloning of the genes encoding an endochitinase from *Phaseolus vulgaris*. Proc Natl Acad Sci USA 1986; 83:6820-4.

- [23] Gaynor JJ. Primary structure of an endochitinase mRNA from Solanum tuberosum. Nucl Acids Res 1988; 16:5210.
- [24] Samac DA, Hironaka CM, Yallay PE, Shah DM. Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana*. Plant physiol 1990; **93**:907-14.
- [25] Shinshi H, Neuhaus JM, Ryals J, Meins F. Structure of a tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequences encoding cysteine-rich domain. Plant Mol Biol 1990; 14:357-68.
- [26] Huang JK, Wen L, Swegle M, Tran HC, Thin TH, Naylor HM, Muthukrishnan S, Reeck CR. Nucleotide sequence of a rice genomic clone that encodes a class I endochitinase. Plant Mol Biol 1991; 16:479-80.
- [27] Newhaus JM, Sticher L, Meins F, Boller T. A short terminal sequence is necessary and sufficient for the targeting of chitinase to the plant vacuole. Proc Natl Acad Sci USA 1992; **88**:10362-6

Received 23-1-1996. Revised 13-6-1996. Accepted 17-6-1996.