Gene Clone and Bioinformatics Analysis of Subtilisin-Like Protease in *Cucumis sativus*

Xue Wang¹, Guangchao Yu¹, Xiangyu Wang¹ and Chunmao Lv²*

¹College of Bioscience and Biotechnology, Shenyang Agricultural University, Shenyang 110866, China.
²College of Food Science, Shenyang Agricultural University, Shenyang 110866, China.

Authors’ contributions
This work was carried out in collaboration among all authors. Author CL designed the study. Author XW analyzed the data and wrote the manuscript. Authors XW and GY conducted the experiments. All authors read and approved the manuscript.

Article Information
DOI: 10.9734/BJI/2019/v23i1300071

Received 22 January 2019
Accepted 09 April 2019
Published 26 April 2019

ABSTRACT

Background: Cucumber target leaf spot (TLS), caused by *Corynespora cassicola* (C. Cassiicola), is a serious disease in cucumber (*Cucumis sativus*) production worldwide. Therefore, cultivating new varieties of TLS resistance of *C. sativus* is an important goal of cucumber breeding. Previous studies have shown that subtilisin-like protease (SUBP) plays an important role in response to *C. Cassiicola* infection in resistant plants.

Objective: In this study, the full-length cDNA of the CsSUBP gene was cloned, and the prokaryotic expression vector was successfully constructed in order to study the effects of subtilisin. Furthermore, vital clues regarding CsSUBP gene involved in TLS resistance of *C. sativus* are gained from the bioinformatics assay.

Method: The CsSUBP gene was identified by sequencing with the intermediate vector pMD18 by designing specific primers and PCR amplification techniques. The prokaryotic expression vector pET30a-CsSUBP was further constructed and identified by colony PCR and EcoR V and Sal I double digestion.

Result: The primary structure of CsSUBP was predicted and analyzed by bioinformatics analysis.

*Corresponding author: E-mail: bt_lcm@126.com, 1997500018@syau.edu.cn;
The results showed that CsSUBP was a weakly acidic protein, N-terminal signal peptide region, including an Inhibitor I9 domain domain. **Conclusion:** The pET30a-CsSUBP prokaryotic expression vector was constructed successfully. This study is convenient for the study of prokaryotic expression and its kinase activity.

**Keywords:** Subtilisin protease; gene clone; bioinformatics assay; prokaryotic expression.

1. INTRODUCTION

Cucumber (**Cucumis sativus**) is one of the largest vegetable crops in China. In recent years, cucumber target leaf spot (TLS), caused by *Corynespora cassiicola* (**C. cassiicola**), is a major disease in cucumber production [1]. At present, the prevention and treatment of TLS are mainly chemical control, chemical pesticide can be used to inhibit the activity of pathogens. However, **C. cassiicola** can easily mutate with environmental changes and develop resistance to various chemical pesticides, so it is difficult to prevent and eliminate the disease efficiently [2]. The best method for controlling the disease in cucumber is the use of resistant cultivars, but few high-resistant varieties of cucumber can enhance resistance of **C. cassiicola** at present. Therefore, understanding the molecular mechanisms and identifying the molecular components involved in the defense responses of the TLS resistant cucumber cultivars will provide a promising approach to restraining this disease.

In our previous study, a kind of subtilisin-like protease (SUBP) from resistant cucumber plants (Jingyou 38) after the fungus inoculation were explored using iTRAQ quantitative proteomics technique. It is speculated that the SUBP may play an important role in cucumber resistance to early invasion of **C. cassiicola**.

SUBP is a serine protease widely found in plants, bacteria, fungi, and parasites [3-8]. Such proteases have a typical Asp/Ser/His catalytic domain in structure. Studies have shown that it is closely related to the stability of the fungal phenotype and the pathogenicity of the pathogen. It plays an important immune stimulating role in the interaction between plants and pathogens [9]. SUBP is also involved in plant cell programmed cell death (PCD) processes [10]. A number of studies have demonstrated the role of SUBP in plant and pathogen interactions. For example, two subtilisin-like proteins, SAS-1 and SAS-2, in oat (**Avena sativa**) exhibit specific activity of mammalian caspase apoptosis protease and participate in the signal transduction of PCD during the non-affinity interaction of oat-Cochliobolus victoriae, thereby inducing HR production [11-12]. Overexpression of Arabidopsis subtilases gene **AtSUBT3.3** activates downstream immune signaling and induces innate immune response in **Arabidopsis** [13].

In this study, the cDNA sequence of CsSUBP gene was cloned and the prokaryotic expression vector was constructed. The primary structure of CsSUBP was predicted and analyzed using bioinformatics analysis methods. On this basis, we can further explore the defense function and molecular mechanisms of SUBP against **C. cassiicola** in cucumber.

2. MATERIALS AND METHODS

2.1 Plant Materials

The cucumber varieties used in the experiments were Jinyou 38 (preliminary experiments showed that this variety was resistant to TLS), which were planted in a greenhouse at 28°C under 16:8 light/dark cycles.

2.2 Full-length cDNA Cloning

Two candidate genes were screened through preliminary test on the basis of the proteomics analysis. The cDNA sequences of **CsSUBP** (A0A0A0K993) from resistant cucumber leaves were clone. The sequence of **CsSUBP** were aligned with cucumber genome database using the service provided byhttp://cucurbitgenomics.org/BLAST. The primers (Table 1) were designed by using the DNAMAN. The PCR fragments amplified from the cDNAs were cloned into TA vectors pMD18-T and were sequenced.

2.3 Construction of Recombinant Prokaryotic Expression Vector pET30a-CsSUBP

Recombinant plasmid pMD18-T-CsSUBP and expression vector pET30a were digested by **EcoR V** and **SalI** respectively. The double-
Table 1. CsSUBP gene primer sequences

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Primer sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CsSUBP-F</td>
<td>GATATCATGTCCGGCCAAATCCACACGCCCTG</td>
</tr>
<tr>
<td>CsSUBP-R</td>
<td>GTCGACCTAATGAAGCCTTGCTGCTCCTCCT</td>
</tr>
</tbody>
</table>

Table 2. Bioinformatics analysis content and tools website

<table>
<thead>
<tr>
<th>Analyze the content</th>
<th>Name of software</th>
<th>Bioinformatics analysis tools website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary structure</td>
<td>ProtParam</td>
<td><a href="http://web.expasy.org/protparam/">http://web.expasy.org/protparam/</a></td>
</tr>
<tr>
<td>Secondary structure</td>
<td>SOPMA</td>
<td><a href="http://npsapbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html">http://npsapbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html</a></td>
</tr>
<tr>
<td>Hydrophobicity</td>
<td>ProtScale</td>
<td><a href="http://web.expasy.org/protscale/">http://web.expasy.org/protscale/</a></td>
</tr>
<tr>
<td>Signal peptide</td>
<td>SignalP 4.0</td>
<td><a href="http://www.cbs.dtu.dk/services/SignalP/">http://www.cbs.dtu.dk/services/SignalP/</a></td>
</tr>
<tr>
<td>Functional domains</td>
<td>SMART</td>
<td><a href="http://smart.embl-heidelberg.de/">http://smart.embl-heidelberg.de/</a></td>
</tr>
</tbody>
</table>

digested DNA fragment had the same sticky ends as the pET30a vector backbone, and the two fragments were ligated into a complete recombinant plasmid using T4 ligase and ligated overnight at 16°C and then trans-formed into E. coli DH5α competent cells, blue and white screening. Recombinant prokaryotic expression vector was identified by direct colony PCR. Recombinant prokaryotic expression vector pET30a-CsSUBP plasmid was extracted and the product was digested by EcoR V and Sal respectively. It was digested for 3 h at 37°C in a constant temperature metal bath and electrophoresed on a 1.5% agarose gel. Plasmids with positive clones were sent for bioinformatics sequencing.

2.4 Bioinformatics Analysis

The specific bioinformatics [14-17] analysis contents and tools used are as follows Table 2.

3. RESULTS

3.1 Amplification of CsSUBP

The reverse transcription of cDNA was used as the template. PCR amplification product was subjected to 1.5% agarose gel electrophoresis and a fragment of approximately 282 bp was obtained as expected (Fig. 1).

3.2 Construction of Prokaryotic Expression Vector

To construct the prokaryotic expression vector, we recycle the small fragments of pMD18-T-CsSUBP (Figs. 2, 3) and large fragments of pET30a plasmid after double digest with EcoR V and Sal. The colony PCR products of recombinant plasmid pET30a-CsSUBP via agarose gel electrophoresis analysis showed 282 bp of the target gene band (Figs. 4, 5).

Sequence results showed that the gene fragments were consistent with the target gene sequence in GenBank (Fig. 6).

3.3 Bioinformatics Analysis of CsSUBP Proteins

Some information about the CsSUBP are obtained by some online tools (shown in table2). The protein molecular weight is 23146.02, chemical formula is C834H1388N282O348S69, isoelectric point (pl) is 5.28. It is weakly acidic and hydrophobic average coefficient is 0.794. It is speculated that CsSUBP may be a hydrophobic protein (shown in Fig 8). Neither signal peptide nor transmembrane region is found in the CsSUBP (shown in Fig 9).

Secondary structure of CsSUBP is analysed by SOPMA. The result showed that CsSUBP consists of 39.78% alpha helix (h), 16.13% extended strand (e), 11.83% beta turn (t) and 32.26% random coil (c) (Fig.7 and Table 3).

A major functional domains of CsSUBP-encoded proteins is found by SMART analysis. The results are shown in Table 4.

Inhibitor_I9 domain, which belongs to a member of the MEROPS family of protease inhibitors I9, a protease pro-peptide inhibitor (sometimes referred to as an activated peptide) that regulates the folding and activity of the proenzyme or zymogen. The anterior segment extends into the enzyme moiety and shields the substrate binding site, thereby promotes enzyme inhibition. Although it often has lower sequence identity, several such propeptides have similar topologies. The propeptide region has an open-loop, sandwich-like anti-parallel α/β, it also has two α-helices and four β-strands with a 2x(β/α/β) topology.
Fig. 1. Electrophoresis profile of PCR products of CsSUBP Gene. A1-A4: CsSUBP PCR amplification products; M: DL2000 DNA Marker

Fig. 2. The colony PCR of recombinant plasmid pMD18-CsSUBP. B1-B3,B5-B8: Colony PCR of CsSUBP gene; B4: water; M: DNA marker DL2000

Fig. 3. Double digestion of recombined vector pMD18-CsSUBP with restriction enzymes.C1-C4: Double digestion products; M: DNA marker DL2000
Fig. 4. The colony PCR of recombinant plasmid of pET30a-CsSUBP. D1-D3, D6: Colony PCR products of CsSUBP gene; D4, D5: Water control; M: DNA marker DL200

Fig. 5. Double digestion of recombined vector pET30a-CsSUBP with restriction enzymes. E1, E2: Double digestion products; M: DNA marker DL2000

Table 3. Prediction the secondary structure of CsSUBP protein

<table>
<thead>
<tr>
<th>Name</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha helix</td>
<td>37</td>
<td>39.78%</td>
</tr>
<tr>
<td>310 helix</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Pi helix</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Beta bridge</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Extended strand</td>
<td>15</td>
<td>16.13%</td>
</tr>
<tr>
<td>Beta turn</td>
<td>11</td>
<td>11.83%</td>
</tr>
<tr>
<td>Bend region</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Random coil</td>
<td>30</td>
<td>32.26%</td>
</tr>
<tr>
<td>Ambiguous states</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Other states</td>
<td>0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
Fig. 6. Sequencing results recombinant vector pET30a-CsSUBP

Fig. 7. The secondary structure of CsSUBP protein. h: Alpha helix; t: Beta turn; e: Extended strand; c: Random coil

Table 4. Functional domain analysis of CsSUBP-encoded proteins

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibitor I9 domain</td>
<td>8(Starting site)-84(Termination site)</td>
<td>4.8e-15</td>
</tr>
</tbody>
</table>
CsSUBP also belongs to the N-terminal propeptide domain of the meropsis family S8A, a subtilisin peptidase. The use of the subtilisin propeptide as a chaperone helps to fold the mature peptidase. The propeptide is removed by proteolytic cleavage to remove the activating enzyme.

4. DISCUSSION

Cucumber is one of the important vegetables in the world, and it is also the largest greenhouse product in China. Cucumber has a cool taste and medicinal value. However, in the process of production, cucumber will be exposed to a variety of diseases such as blight, powdery mildew, and target leaf spot. Among them, the incidence of target leaf spot is extremely high, which causing serious economic losses [18]. Studies have shown that breeding and using disease-resistant varieties are the most direct, effective, and economical measures for disease control.
Our iTRAQ analysis on the inoculation of cucumber leaves showed the expression level of a subtilisin-like protease (A0A0A0K993) was significantly decreased at 6 h after pathogen infection. In addition, quantitative real-time PCR technology found that its corresponding genes expression were also decreased a downward trend after inoculation of pathogen. We speculated that it may be involved in the early invasion process of cucumber against C. cassicola.

Subtilisin-like protease (SUBP) is a serine protease, which can regulate plant embryogenesis and organ development, fruit ripening and senescence [19-22]. It also plays an important role in the interaction between plant and pathogen [9]. Plant-pathogen interaction begins with the pathogen contact host and ends with a significant disease or susceptibility response [23]. When plant was attacked by pathogen, it will produce two immune defense lines (MTI and ETI). Studies have shown that in plants such as tomato [24], soybean [25], and wheat [26], the subtilisin-like protease can regulate its immune response and participate in intercellular signal transduction during interaction. However, whether it will stimulate the immune response in cucumber is still unclear. In this study, bioinformatic analysis of the subtilisin-like protease showed that it contained a propeptide domain (Inhibitor_I9 domain), which may be the main disease resistance domain of the protein.

E. coli prokaryotic expression system has the characteristics of simple operation, short culture period and high expression level, which is beneficial to the study of protein structure and function [27]. In addition, the prokaryotic expression vector pET-30a has the advantages of strict regulation and exquisite design of multiple cloning sites. In this research, the SUBP gene was cloned from Jinyou 38 and the prokaryotic expression vector pET30a-CsSUBP was constructed to facilitate the establishment of CsSUBP-encoded recombinant protein. It also lays the foundation for studying the role of protein function and kinase activity in plant disease resistance.

5. CONCLUSION

We have successfully amplified and purified the coding sequence of CsSUBP gene in cucumber, and constructed the recombinant prokaryotic expression vector pET30a-CsSUBP. Bioinformatics analysis shows that CsSUBP belongs to the hydrophobic protein and has no signal peptide at the N-terminus. It contains a functional domain. Its protein chemical formula is C834H1388N282O348S69, an isoelectric point (PI) is 5.28 and a molecular weight is 23146.02. Our study is a preparation for revealing the functions and mechanisms of CsSUBP against C. cassicola in cucumber.

ACKNOWLEDGEMENTS

This work was supported by Liaoning Provincial Natural Science Fund Guidance Program20180550919.

COMPETING INTERESTS

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

7. Mizutani O, Shiina M, Yoshimi A, et al. Substantial decrease in cell wall α-1, 3-glucan caused by disruption of the kexB


© 2019 Wang et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history: The peer review history for this paper can be accessed here: http://www.sdiarticle3.com/review-history/48390