The Elucidation on the Involvement of Indica Rice Oryzacystatin-I Gene in Response to Abiotic Stresses via Bioinformatics Analysis

Nor Mustaiqazah Juri¹,², Noor Azmi Shaharuddin¹, Norliza Abu Bakar²
¹Department of Biochemistry, Faculty of Biotechnology and Molecular Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
²Malaysian Agricultural Research and Development Institute, Persiaran MARDI-UPM, Ibu Pejabat MARDI, 43400 Serdang, Selangor, Malaysia

Abstract: Compared with other plant species, the study on the potential of rice cystatin gene (Oryzacystatin-I) in response to abiotic stress is underexplored. Regarding this, isolation and sequence analysis of OCI gene was conducted on three Indica rice varieties, named as MR219, MR211 and AERON 1. Unique features of phytocystatins family were found on the isolated sequences, including GG residue, LARFAVTHEN, QVVAG and PW motif. Phylogenetic analysis indicates that this gene is a putative orthologous for corn, wheat and barley cystatin family. Various cis-acting regulatory elements that associated with stress responses were identified at the upstream regions of this rice cystatin gene, indicating the potential of this gene to be used in combating the abiotic stress issues in world rice production.

Keywords: Abiotic stress, Indica rice, Oryzacystatin-I

Date of Submission: 14-09-2018
Date of acceptance: 17-10-2018

I. Introduction

Since 1990, Oryzacystatin-I (OCI) had been extensively studied to combat the destructive effect of biotic stress in crops. It was found to work as a potent inhibitor against insects [1, 2], nematodes [3] and viruses [4]. This gene was also had been transformed and overexpressed into several crops including potato [5], eggplant [6] and sweet potato [7]. Interestingly, OCI had been engineered into OC-IΔ086 and exogenously expressed in tomato, forming a more potent inhibitor to Globodera pallida compared to the wild OCI [8]. However, compared to other plant cystatin, the study on the role of OCI in responses to abiotic stresses was considered as far left behind, where the accumulation of cystatin messages in other plant species induced by abiotic stress treatments were progressively reported in other plant species including in barley [9], cowpea [10], chestnut [11], Arabidopsis thaliana [12] and Jatropha curcas [13]. These reveal its bi-functional potential to be utilized in combating the biotic and abiotic stress dilemma in rice production itself [14].

In this study, by using a specific primer design, OCI coding sequence were isolated from three Indica rice varieties, named as MR219, MR211 and AERON 1. Besides the general structural characterization, in silico functional prediction of OCI gene was also been conducted through identification of signal peptide translocation and cis-acting elements in 5’ regulatory regions of OCI. Each gene in eukaryotic system possess a unique combination of cis-acting elements in their promoter region, specifically along 1500 base pair (bp) upstream of the gene transcriptional start site. Those cis-acting elements are important transcriptional gene regulatory units as they determine the temporal and spatial expression of a gene in respond to various biological and stress conditions [15]. To date, there is no analysis on OCI cis-acting elements has been conducted. Besides, analysis of phylogenetic tree generated allowed us to identify paralogous genes of OCI.

II. Methodology

2.1 Plant Material

Seeds of MR219, MR211 and AERON 1 rice varieties were were supplied by Malaysian Agricultural Research and Development Institute (MARDI) Seberang Perai, Pulau Pinang. The seeds were surface-sterilized by soaking them in 70% of (v/v) Clorox® solution (5% sodium hypochlorite) and were let to germinate on wetted cotton at 28°C in the dark for a week before transferred into soil at the glass house. The water level were maintained at the 3 cm above the soil surface and fertilized with NPK and urea for every three weeks. At two weeks of flowering stage, the immature rice seeds were harvested and grinded in liquid nitrogen into fine powder. Approximately 100 mg of grinded samples were transferred into pre-chilled 2 ml centrifuge tubes and proceed for total RNA extraction.
2.2 Total RNA extraction and cDNA synthesis

Total RNA was extracted with RNeasy Plant Mini Kit (Qiagen, USA) according to manufacturer’s protocol. The quality of extracted RNA was analyzed in 1% agarose through gel electrophoresis and was quantified with NanoDrop ND 1000 spectrophotometer (Thermo Scientific, USA). cDNA were synthesized by using iScript cDNA Synthesis Kit (Bio-Rad, USA) following to manufacturer’s protocol.

2.3 Primer design and polymerase chain reaction (PCR)

A set of specific primer for OCI gene amplification were design manually based on OCI gene sequence of Nipponbare’s rice variety, that is available in GenBank of NCBI website (Accession number NM_001051085). The sequences of the forward and reverse primer are 5’-ATGCCGAAAATATCGAGTCGC-3’ and 5’-TTAGGCAATTGCACCTGGCAT-3’, respectively. A final volume of 20 µl PCR reaction mix were prepared, containing 0.5 µg of cDNA template, 0.5 mM of each forward and reverse primer and GoTaq® Flexi DNA Polymerase (Promega, USA) according to the manufacturer’s protocol. PCR was conducted by TECHNE-TC 521 thermal cycler (TECHNE, UK) with the following parameters: initial denaturation at 94°C for 3 minutes, followed by 20 cycles of 94°C for 45 seconds (denaturation), 59°C at 45 seconds (annealing), 72°C for 1 minute (extension) and 72°C for 6 minutes. 3 µl of the PCR products were run on 1% agarose gel electrophoresis and the size of band was determined by VC 100bp DNA Ladder (Vivantis, USA).

2.4 Molecular cloning of Oryzacystatin-I gene

PCR product with expected size of Oryzacystatin-I gene (423 bp) was purified using QIAquick PCR Purification kit (Qiagen, USA) and ligated into pGEM-T easy vector (Promega, USA) by overnight incubation at 4°C. Blue/white colony screening was conducted by preparing Luria-Bertani (LB) agar supplied with 10 µg/µl of ampicillin, spread with 40 µg/µl X-gal and 40 µl of 500 mM IPTG. The ligation mixture was spread on the agar and cultured for overnight. 3 white colonies were selected randomly and colony PCR was carried out using the same parameters of previous PCR for recombinant colony confirmation. The confirmed recombinant colony was cultured into LB broth supplied with 10 µg/µl of ampicillin for overnight. Recombinant plasmid was extracted by using GeneJET Plasmid Miniprep Kit (Thermo Scientific, USA) and was send to 1st Base (Malaysia) for DNA sequencing.

2.5 Sequence and structural analysis

Multiple sequence alignment of the sequencing result and their putative protein sequences were conducted by using CLC Main Workbench version 7 program (http://www.clcbio.com) with default parameter. For plant cystatins domain organization check, the sequences were subjected to BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) databases. The signal peptide translocation was detected using SignalP 3.0 program (http://www.cbs.dtu.dk/services/signalP), whereas putative three-dimensional structure was generated by SWISS-MODEL program.

2.6 Phylogenetic tree construction

Unrooted dendrogram of the predicted protein sequences with other plant species cystatin were constructed using Molecular Evolutionary Genetics Analysis (MEGA) version 6.0 software (Kumar et al. 2013). Bootstrap analysis was conducted with 1000 replication through neighbour-joining method.

2.7 Identification and analysis of putative cis-acting elements

The sequencing results were subjected for nucleotide blast search against Oryza sativa (indica cultivar-group) genome present in NCBI database to determine the location of transcriptional start site of oryzacystatin-I in rice chromosome. Then sequence at 1500 bp upstream of the transcriptional start site, or known as 5’ regulatory region were extracted and scanned in PlantCARE database [16] for the identification of putative cis-elements of OCI gene.

III. Results

3.1 Cloning and nucleotide sequence analysis of OCI

By comparing with the Nipponbare’s OCI coding sequence (CDS) in NCBI database (Accession number NM_001051085.1), the sequencing result indicates that complete CDS of OCI gene with the size of 423 bp were successfully isolated from the all three rice varieties. The deposited sequences of OCI from AERON 1, MR219 and MR211 in NCBI databases can be referred with accession number of KM983023.1, KM983024.1 and KM983025.1 respectively. Nucleotide BLAST of Nipponbare’s OCI CDS revealed that they had shared 99% sequence identity (Table 1), which each of them encoded for 140 amino acids residues.
The Elucidation on the Involvement of Indica Rice Oryzacystatin-I Gene in Response to Abiotic ...

Table 1. Nucleotide BLAST result of MR219, MR211 and AERON 1 OCI sequences using BLASTN. CDS of Nipponbare’s OCI was used as the query.

<table>
<thead>
<tr>
<th>Description</th>
<th>Max score</th>
<th>Description</th>
<th>Max score</th>
<th>Description</th>
<th>Max score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AERON 1</td>
<td>776</td>
<td>MR219</td>
<td>771</td>
<td>MR211</td>
<td>771</td>
<td></td>
</tr>
</tbody>
</table>

Through BLASTP analysis of their deduced protein sequence, two conserved motifs of cystatin superfamily were identified. They are a Glycine residue located near the N-terminal and a QXVXG motif (Fig. 1A). Besides, signal peptide were also been identified for the all three sequences (Fig. 1B). In order to identify the specific domain structure for plant cystatin, their deduced amino acid was aligned and the detected motifs were marked.

![Fig. 1. Sequence analysis of deduced OCI amino acid sequences of AERON 1, MR 219 and MR 211. (A) Graphical summary of BLASTP result had identified conserved domain of cystatin superfamily in all three sequences. (B) Signal peptide was detected in all three OCI sequences of AERON 1, 219 and MR 211. Analyses were conducted using SignalP software.](image)

3.2 Structural Analysis of deduced protein sequence of OCI

In order to identify any amino acid differences between isolated OCI sequences of MR 219, MR 211 and AERON 1, their deduced protein sequences were aligned together with Nipponbare’s OCI sequence. Only single amino acid difference was detected at 25th amino acid residue of MR219 OCI, where Arg25 that being conserved in other sequences was changed with Glu25 (Fig. 2A). All the motif signatures (GG, QVVAG and PW) that responsible for their inhibitory properties were conserved along the 140 amino acid nucleotides without any changes.

The predicted three-dimensional structures of AERON 1, MR219 and MR211 OCI were established by using known crystal structure of OCI as template, as suggested by the program. As being expected, no any structural variation was observed due to their high sequence identity to each other. Each of them conserving an α-helix that spanning along the LARFAV motif and five β-sheets (β1, β2, β3, β4, β5) as indicated by arrow sign in Fig. 2A. Each motifs of QVVAG and tryptophan (W) forming a loops as indicated by asterisks (*) and crosses (+) in Fig. 2B, where this site will had a direct interaction of their target proteinases.
The Elucidation on the Involvement of Indica Rice Oryzacystatin-I Gene in Response to Abiotic ....

3.3 Phylogenetic tree and sequence analysis

The branching pattern of the tree showed that all plant cystatins were well-clustered into four main branches, thus divided them into four groups (Group 1, 2, 3 and 4). Although two of the main branch have low bootstrap values on their nodes (Group 1: 42% and Group 3: 36%), their distribution pattern was still align with previous tree, constructed by Martinez [17] and Christoff and Margis [18]. All plant cystatin were distributed all over within group 1 to 3, except for broccoli cystatin (BoCPI) where they had form their own groups, indicated that broccoli cystatins were closely related to each other compared to such genes from other plant species.

Except for group 2 and 4, the nodes of the major branch showed low bootstrap score (around 40%). However their nodes for most of minor groups showed high bootstrap scores, with the value of >90%. As being expected, all deduced OCI that had been isolated from MR 219, MR 211 and AERON 1 were clustered together with Nipponbare's OCI at the same branch.

Since OCI is grouped under Group 1, further discussion on this tree will be focused on this group only. Based on the branching pattern, ten orthologous groups were predicted within this group. They are wheat/barley (WC1/WC3/HvCPI-2), barley/rice/corn/sugarcane (HvCPI-1/OCI/CC3/CaneCPI-1), corn/rice (CC1/CC2/OCI), Arabidopsis/Brassica rapa (AtCYS-1/BrCYS-3/BrCYS1-2), CC5/OCIII, AtCYS-7/BrCYS7), Brassica rapal/Arabidopsis (BrCYS3-1/AtCYS-3), apple/Amaranthus (MdCYS18/MdCYS21/AhCPI), Arabidopsis/Cakile maritima (AtCYS-6/CmC) and corn/barley/rice (CC4/HvCPI-4/OCXII).
The Elucidation on the Involvement of Indica Rice Oryzacystatin-I Gene in Response to Abiotic...

Fig. 3. Unrooted phylogenetic tree built by using neighbour-joining method with 1000 bootstrap replication size. The three predicted protein of the isolated OCI sequences (MR219, MR211 and AERON 1) were combined with whole phytocystatin gene family from Nipponbare’s rice variety (OCI-XII), barley (HvCPII-13), corn (CC1-10), Cakile maritima (CmC), Amaranthus hypochondriacus (AhCPI), wheat (WC1-5), sugarcane (CaneCPI-1), Arabidopsis (AtCYS1-7), broccoli (BoCPI-5) and several selected phytocystatin family members from apple (MdCYS) and Brassica rapa (BrCYS).

3.4 Analysis of OCI cis-acting elements

Nucleotide blast (BLASTN) of isolated OCI from MR 219, MR 211 and AERON 1 rice against Oryza sativa (Indica) Refseq-genomic in NCBI database showed that transcriptional start site (ATG) of OCI in chromosome 1 are located at the base number 35,789,133, thus the 1500 bp upstream of the transcription start site were located starting from base number 35,787,632 to 35,789,132. Scanning of this region enabled the prediction of various group of cis-elements that are possibly involved in the regulation and expression of OCI, which are related to plant hormone responses, environmental responses, light responses and involved in specific organ and growth stages. The details of the listed cis-elements present at the 5’ regulatory region and their predicted function can be referred in Table 2.

Various light responsive elements such as ACE, Box 4, Box I, G-Box, GT1-motif, I-box and Sp1 were distributed with different frequencies along the OCI promoter region, with G-Box as the most obvious compared to others. There are many cis-acting elements that are function as the transcription activator in plant response to various abiotic factors. The cis-elements are W-box, MBS, ARE and GC-motif, and LTR where each of them are respectively responsive elements for wounding, drought, oxygen depletion and cold.

Motifs associated with seed-specific regulation such as Skn-1 and O2-site motifs were abundantly found on OCI promoter region. Spreading pattern of most cis-elements in the 5’ regulatory region that are far from the translational start site (ATG) indicates that OCI gene expression may be well regulated at the distal region.

DOI: 10.9790/264X-0405023036 www.iosrjournals.org 34 | Page
Table 2. Potential cis-acting regulatory elements identified in the 5′ regulatory sequences of rice (Oryza sativa Indica cultivar-group) of OCI gene. The 1.5 kb of 5′ regulatory region was analyzed using Plant CARE databases.

<table>
<thead>
<tr>
<th>Types</th>
<th>Cis-elements</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abiotic Acid</td>
<td>ABRE</td>
<td>Cis-acting element involved in the abiotic acid responsiveness</td>
</tr>
<tr>
<td>Auxin</td>
<td>ARF (AnRE)</td>
<td>Cis-acting regulatory element involved in Auxin responsiveness</td>
</tr>
<tr>
<td>Methyl Jasmonates</td>
<td>MeJA motif</td>
<td>Cis-acting regulatory element involved in the methyl jasmonates responses</td>
</tr>
<tr>
<td>Salicylic acid</td>
<td>TCA-element</td>
<td>Cis-acting element involved in salicylic acid responsiveness</td>
</tr>
<tr>
<td>Light</td>
<td>ACE</td>
<td>Cis-acting element involved in light responsiveness</td>
</tr>
<tr>
<td>Wound</td>
<td>W-box</td>
<td>Fungal elicitor responsive element</td>
</tr>
<tr>
<td>Drought</td>
<td>MBS</td>
<td>MYB binding site involved in drought-inducibility</td>
</tr>
<tr>
<td>Cold</td>
<td>LTR</td>
<td>Cis-acting element involved in low-temperature responsiveness</td>
</tr>
<tr>
<td>Oxygen deprivation</td>
<td>ABE</td>
<td>Cis-acting regulatory element essential for the anaerobic induction</td>
</tr>
<tr>
<td>Seed</td>
<td>Skn-1 motif</td>
<td>Cis-acting regulatory element required for endosperm expression</td>
</tr>
<tr>
<td></td>
<td>O2-site</td>
<td>Cis-acting regulatory element involved in zein metabolism regulation</td>
</tr>
</tbody>
</table>

IV. Discussions

There are dozens of plant cystatin transcript profiling studies have been reported in many plant species in responses to various abiotic stresses [12,13]. Realizing on its potential to be utilized to combat the abiotic stress dilemma in world rice productions, we are trying to re-dissect on this gene by using local Indica Malaysian rice (MR211 and MR219) and an aerobic rice released by International Rice Research Institute (IRRI) known as AERON 1. The availability of computational tools for nucleotide sequence analysis, plus the dozens of new reports on plant cystatins has become the value added for this research.

Little is known about the essential transcriptional regulatory components that contribute to any particular pattern of plant cystatin transcript profiling. Being a part of transcriptional gene regulatory units, cis-acting elements and transcriptional factors control multiple stress responses and biological processes [19]. Taking this into account, we had characterized the localization and abundancy of important cis-acting elements located across OCI’s 5′ regulatory region. It is not shocking to see the high frequency of seed-specific regulation cis-elements on OCI promoter region, underlying the dominant role of OCI in seed. It is well-reported that OCI transcript was highest expressed in the rice seeds, specifically two weeks after flowering period [20], where OCI is very crucial to protect the degradation of immaturesd storage protein the seeds by papain-proteinases.

As being reported by Christoff and Margis [18], the expression of OCI was highly up regulated when the plants were infected with blast fungus M. Grisea and treated with high salinity, abiotic (ABA) and chitosan. At this point, the finding of fungal elicitor responsive element (W-box) and ABA responsive elements (ABRE, motif IIb and CE1) showed the correlation of this finding with their result.

For salt responses, even though there is no salt responsive element being scanned on OCI promoter such as DRE motif, but the present of ABA responsive elements such as ABRE, motif IIb and CE1 might be the alternative elements for salt responses in OCI. As stated by Suh [21], ABA responsive elements were important to regulate responses to abiotic factor such as drought and salinity.

The detection of cis-element that involved in drought-inducibility (MBS) on the promoter region of OCI showed a high probability in their involvement during drought condition. As been reported by Subburaj [22], several cystatin gene family from Brachypodium distachyon L. had showed an up regulation of expression in response to drought stress. In other study, an over-expression of cystatin gene from Malus prunifolia in Arabidopsis had enhanced the plant tolerance to drought stress [23].
Based on the correlation present between the orthologue group and cis-elements present in gene promoter region, it is hypothesized that orthologue gene might contain high similarity cis-elements in their promoter region. However, to prove this, the promoter region of CC1 and CC2 need to be sequenced first.

V. Conclusion

Complete CDS of OCI genes were successfully isolated and sequenced from three Indica rice; MR219, MR211 and AERON 1. The bioinformatics analysis on the sequenced OCI had revealed various cis-acting regulatory elements that involved in environmental stress which including of wound, drought, cold and oxygen depletion were identified at their 5′ regulatory sequences. More than that, there were present of cis-acting regulatory elements that involved in plant hormone and light responses. The identification of those regulatory elements will not only provide a full understanding of OCI transcriptional system, but also can be used to guide future OC experimental work, especially those related to abiotic stress responses.

Acknowledgements

This project was funded by Fundamental Research Grant Scheme (FRGS). We would like to acknowledge Universiti Putra Malaysia (UPM) Serdang and Malaysian Agricultural Research and Development Institute (MARDI) for providing infrastructure and lab facilities.

References