C. R. Biologies 337 (2014) 683–690

Contents lists available at ScienceDirect

Comptes Rendus Biologies

www.sciencedirect.com

Molecular biology and genetics/Biologie et génétique moléculaires

Characterization of *P5CS* gene in *Calotropis procera* plant from the *de novo* assembled transcriptome contigs of the high-throughput sequencing dataset

Ahmed M. Ramadan [a,b,*], Sameh E. Hassanein [c,d]

[a] *Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), PO Box 80141, Jeddah 21589, Saudi Arabia*
[b] *Plant Molecular Biology Department, Agricultural Genetic Engineering Research Institute (AGERI), Agriculture Research Center (ARC), Giza, Egypt*
[c] *Bioinformatics and Computer Networks Department, Agricultural Genetic Engineering Research Institute (AGERI), Agriculture Research Center (ARC), Giza, Egypt*
[d] *College of Biotechnology, Misr University for Science and Technology (MUST), 6th October city, Egypt*

ARTICLE INFO

Article history:
Received 16 July 2014
Accepted after revision 13 September 2014
Available online 7 October 2014

Keywords:
P5CS
3D modeling
Calotropis procera
Abiotic stress

ABSTRACT

The wild plant known as *Calotropis procera* is important in medicine, industry and ornamental fields. Due to spread in areas that suffer from environmental stress, it has a large number of tolerance genes to environmental stress such as drought and salinity. Proline is one of the most compatible solutes that accumulate widely in plants to tolerate unfavorable environmental conditions. Plant proline synthesis depends on Δ-pyrroline-5-carboxylate synthase (*P5CS*) gene. But information about this gene in *C. procera* is unavailable. In this study, we uncovered and characterized *P5CS* (P5CS, NCBI accession no. KJ020750) gene in this medicinal plant from the *de novo* assembled transcriptome contigs of the high-throughput sequencing dataset. A number of GenBank accessions for P5CS sequences were blasted with the recovered *de novo* assembled contigs. Homology modeling of the deduced amino acids (NCBI accession No. AHM25913) was further carried out using Swiss-Model, accessible via the EXPASY. Superimposition of *C. procera* P5CS-like full sequence model on *Homo sapiens* (P5CS_HUMAN, UniProt protein accession no. P54886) was constructed using RasMol and Deep-View programs. The functional domains of the novel P5CS amino acids sequence were identified from the NCBI conserved domain database (CDD) that provide insights into sequence structure/function relationships, as well as domain models imported from a number of external source databases (Pfam, SMART, COG, PRK, TIGRFAM).

1. Introduction

Environmental stress causes non-desirable effects on plants' growth and productivity, especially drought and salinity [1]. Synthesizing and accumulating compatible osmolytes in plants, such as proline and glycine betaine, facilitate coping with this condition [2,3].

Amino acid proline is an α-amino acid, and is not an essential amino acid, which means that living organisms can synthesize it. It is unique among amino acids, because it contains a secondary amino group. In addition to its role in protein forming, proline is one of the most widely distributed compatible solutes that accumulate in plants and bacteria during unfavorable environmental conditions

* Corresponding author. Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), PO Box 80141, Jeddah 21589, Saudi Arabia.
E-mail address: ahmedramadan782@yahoo.com (A.M. Ramadan).

http://dx.doi.org/10.1016/j.crvi.2014.09.002
1631-0691/

[4,5]. The role of proline in the endurance of the environmental stress is still a matter of intensive research [6–8].

In plants, the synthesis of proline depends on two different precursors, glutamate and ornithine, through two different cycles [9,10]. In the first cycle, proline is produced *via* two reduction reactions of glutamate in which two enzymes catalyze these reactions, e.g., Δ-pyrroline-5-carboxylate synthase (P5CS) and pyrroline-5-carboxylate reductase (P5CR). P5CS is an enzyme activating glutamate through the phosphorylation process. This enzyme also reduces the product to form glutamate semi-aldehyde (GSA) [11]. In the second cycle of proline synthesis, ornithine turns to form pyrroline-5-carboxylate with the catalysis of orn-Δ-aminotransferase (OAT). This enzyme exists in mitochondria [7]. However, when plants experience adverse environmental conditions, proline is synthesized mainly through the first cycle. This has been demonstrated through analyses of the expression of P5CS and P5CR in *Arabidopsis thaliana* and moth bean plants [11–13].

Calotropis procera (*C. procera*) is a drought-tolerant wild plant. It belongs to the *Asclepiadaceae* family and is characterized as a sustainable evergreen toxic shrub. Seed spreads mainly by wind and can be transmitted by animals as well. Therefore, this plant is seen along roadsides, and the edges of lakes and native pastures, while scattered in desert areas [14,15]. *C. procera* is native to west and east Africa, and south Asia, while naturalized in Australia, Central and South America, and the Caribbean islands [15–17]. It provides an excellent source of genes for drought and salt tolerance. In previous work, we found that proline is increased in this plant when irrigated [18]. This finding is contrary to the conclusions of most researchers [2,3]. We suggest that this plant might need proline in another pathway under temporary irrigation. However, the biological significance of P5CS in *C. procera* has not been described. Increasing information about plant genomes in conjunction with bioinformatics tools and databases has led to the availability of new insights into the study of different genes that may be keys to stress responses in plant [19,20].

In this study, we uncovered and characterized one *P5CS*-like gene in this medicinal plant from the *de novo* assembled transcriptome contigs of a high-throughput sequencing dataset. We also compared the sequence as well as the three-dimensional (3D) structure of the obtained P5CS-like protein with those of other plant species.

2. Materials and methods

2.1. Sample collection and isolation of total RNA

Three leaf discs of *C. procera* were collected from Jeddah region (KSA, latitude 21°26′6.00, longitude 39°28′3.00 in September 2012 (with temperature of 37 °C, and air humidity of 70–75%). The samples were frozen in liquid nitrogen (50 mg tissue each) and total RNA extraction was performed using RNeasy Plant Mini Kit (Qiagen, cat. No. 74903). To remove DNA contaminants, 3 μL of 10 mg/mL RNase A, DNase and protease-free Thermo Scientific cat No. EN0531) were added to the RNA samples, and the tube was incubated at 30 °C for 15 min. The RNA concentration in different samples was estimated by measuring the optical density at 260 nm according to the equation: RNA concentration (μg/mL) = OD260 × 40 × dilution factor. RNA samples were sent to Beijing Genomics Institute (BGI), Shenzhen, China, for deep sequencing, and dataset were provided for analysis.

2.2. NGS sequence

Whole-RNA-seq, paired-end short-sequence reads of *C. procera* were generated using the Illumina Genome AnalyserIIx (GAIIx) according to the manufacturer's instructions (Illumina, San Diego, CA).

2.3. Sequence filtering and bioinformatics analysis

The raw sequencing data were obtained using the Illumina python pipeline v. 1.3. For the obtained libraries, only high-quality reads (quality > 20) were retained. Then, a *de novo* assembly of the obtained short (paired-end) read dataset was performed using assembler trinityrna-seq_r20131110 [21] followed by the creation of putative unique transcripts (PUTs) with a combination of different *k*-mer lengths and expected coverage.

Twenty *P5CS* sequences (Table 1) belonging to other plant species were obtained from GenBank and used as a reference for blasting (http://www.ncbi.nlm.nih.gov/BLAST) our obtained library (the yielded EST assemblies from Velvet program) to identify contigs with CpP5CS-like sequence.

Assemblies were mapped to *Apocynum venetum* accession number EF160132 using SAOP [22]. The number of reads aligned was 6577, with an average coverage of 327.33 and the length of consensus sequence, including *C. procera* P5CS-like (*CpP5CS*-like) gene, equals 2154 nt (Fig. 1).

2.4. Determination of phylogenetic relationships

The maximum-likelihood method [23] was used to build a dendrogram and CLC Genomics Workbench was used to allow doing bootstrap analysis. A bootstrap value is attached to each branch to indicate the confidence level in this branch.

2.5. The 3D homology modeling

Homology modeling was carried out using Swiss-Model, a protein-modeling server, accessible *via* the EXPASY (http://www.expasy.org/). Superimposition of *Cp*P5CS-like amino acid sequence model on those of other P5CS proteins was constructed using RasMol (http://www.umass.edu/microbio/rasmol/), and Deep-View programs (http://spdbv.vital-it.ch/). The functional domains were identified from the NCBI's conserved domain database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), which uses 3D structure information to