Egyptian Journal Of

GENETICS AND CYTOLOGY

INTERNATIONAL JOURNAL DEVOTED TO GENETICAL AND CYTOLOGICAL SCIENCES

Published by

THE EGYPTIAN SOCIETY OF GENETICS

Volume 45 January 2016 No. 1

CHARACTERIZATION OF PHOTOSYSTEM TRANSMEMBRANE GENES UNDER SUDDEN WATER SUPPLY IN *Calotropis procera*

HALA F. EISSA[1,2]

1. *Agricultural Genetic Engineering Research Institute (AGERI), Agric. Res. Center (ARC), Egypt*
2. *College of Biotechnology, Misr University for Science and Technology (MUST), Egypt*

Scarcity of water or drought is among the environmental constraints that affect crop growth and productivity worldwide (Chaves *et al*., 2003). It has been estimated that around 45% of the world's agricultural lands are subject to drought (Bot *et al*., 2000). Plants exhibiting water deficit respond rapidly at the physiological, cellular and molecular levels (Chaves *et al*., 2009). Understanding the molecular mechanisms by which plants respond to drought stress can help in the development of drought-tolerant crops *via* molecular breeding and transgenesis. The implementation of the available genomic and transcriptomic databases with the high-throughput bioinformatics tools can increase the chance to uncover target genes and detect their contribution towards drought stress tolerance (Mochida and Shinozaki, 2010).

Photosynthetic capacity is often decreased when plants are subjected to environmental stresses such as drought (Lawlor and Cornic, 2002), salt (De Souza *et al*., 2005) or heat (Sainz *et al*., 2010). It was proposed that the decrease is due to the stomatal closure (Loreto *et al*., 2003) and/or metabolic constraints such as reduction in the activity of ribulose-1,5-bisphosphate carboxylase/oxygenase (Loreto *et al*., 2003) as well as in the plant's ability to synthesize ATP (Tezara *et al*., 2008) and transport electrons during photosystem I (PSI) and photosystem II (PSII) (Hao *et al*., 1999). PSI and PSII are two multimeric chlorophyll-binding protein complexes embedded in the thylakoid membranes (Chitnis, 2001; Barber, 2002). In the photochemical reaction, PSII oxidizes water to produce molecular dioxygen, reduces the plastoquinone, which undergoes reduction by the cyt b_6f complex and donates electrons to the oxidized reaction center of PSI, namely $P700^+$ (Kargul and Barber, 2008). The photo-activated PSI uses reducing equivalents derived from PSII to reduce the ferrodoxin and convert $NADP^+$ to NADPH (Rochaix, 2011). In this way, electrons flow from PSII to PSI *via* the cyt b_6f complex leading to the formation of electrochemical potential gradient across the thylakoid membrane powering the activity of ATP synthase to covert ADP to ATP (Kargul and Barber, 2008). Subsequently, the produced NADPH and ATP are used for CO_2 assimilation in the photosynthetic dark reactions of the Calvin–Benson cycle (Rochaix, 2011).

The photosystem I consists of two functional moieties, the reaction center and the light-harvesting complex (Busch and Hippler, 2011). The central part of the reaction center is formed by the heterodimer of two large transmembrane protein subunits PsaA and PsaB (Nelson and Ben-Shem, 2005). The reaction center is attached to chlorophyll molecules that serve as antenna to capture light energy (Amunts *et al*., 2007). The C termini of both subunits are located on the lumenal side, whereas the N termini are located in the stroma side. The membrane integral parts of PsaA to PsaB share similarities in their amino acid sequence forming pseudo-symmetry (Amunts and Nelson, 2009). The reaction center of PSII core consists of two homologous proteins (D1 and D2), which have five transmembrane α-helices each (Barber, 1987). Two Chl (chlorophyll)-containing proteins (CPs), namely CP43 and CP47, associate closely to the D1 and D2 proteins. The CP 47 comprises six transmemebrane helices arranged in a circular manner (Barber, 2002). The ability of PSI and PSII systems to capture and convert sunlight energy is highly dependent on the precise arrangement of their protein subunits (Amunts *et al*., 2010). Therefore, a detailed knowledge of the three-dimensional structure is required in order to understand these mechanisms at the molecular level.

Desert plants provide a good reference to the adjustment in their photosynthetic system in response to drought stress (Xu *et al*., 2010). The photoprotective mechanisms involve photosynthetic parameters adjustments in photochemistry, leaf morphology and anatomy (Havaux and Niyogi, 1999). Boutraa (2010) found that *Calotropis procera* had a high photosynthesis capacity throughout the year even during the dry season. *C. procera* is a desert evergreen shrub, which grows in arid and semi-arid regions. The author and previous reports (ex., Orwa *et al*., 2009) suggested the presence of particular strategies of drought tolerance in this plant. By the availability of sudden limited amount of water, Ramadan *et al*. (2014) found that *C. procera* has developed a capacity to respond fast by remodeling its metabolic machinery towards growth, through the increase in biosynthesis of essential amino acids and the increase in structural lipids of the photosynthetic membranes.

The objectives of the present study are: 1) characterization of photosystem I P700 chlorophyll a apoprotein A1 (PsaA), photosystem I P700 chlorophyll a apoprotein A2 (PsaB), photosystem II protein D1 (PsbA) and photosystem II CP47 protein (PsbB) genes and encoded proteins in *C. procera* from the *de novo* assembled transcriptome contigs of high-throughput sequencing dataset, 2) comparing the three-dimensional (3D) structure(s) of the obtained PsaA, PsaB, PsbA and PsbB deduced amino acid sequences to other plant species and 3) evaluate the transcript levels for the studied genes after sudden supply of a limited amount of water to this desert-grown plant using quantitative real-time RT-PCR (qPCR).

MATERIALS AND METHODS

Sample collection and isolation of total RNA

Changes in water availability experiment was performed as reported by Ramadan *et al*. (2014). In this experiment, samples were taken from three similar-sized *C. procera* (Taxonomy ID: 141467) plants. Control samples were taken 1 hour post-dawn, at midday and 1 hour pre-dusk. Next day, each plant received 25 L dH_2O and samples were collected 2 h (1 hour post-dawn), 6 h (at mid-day) and 12 h after watering (1 hour pre-dusk) and immediately frozen in liquid nitrogen. Total RNAs were isolated using RNeasy Plant mini kit (Qiagen, cat. No. 74904). DNA contaminants were removed using IRQ1 RNase-Free DNase (Promega, cat. no. M6101). Concentration of RNA was determined using a NanoDrop spectrophotometer.

RNA sequencing and analysis

RNA-Seq to recover paired-end short sequence reads of *C. procera* was performed at Research Technology Support Facility, Michigan State University, East Lansing, USA using the Illumina Miseq according to the manufacturer's instructions (Illumina, San Diego, CA). The raw sequencing data were obtained using the Illumina python pipeline v. 1.3. For the obtained libraries, only high-quality reads (quality > 30) were retained. Assembler Trinityrna- seq_r20131110 (Haas *et al*., 2013) was used to perform a *de novo* assembly of the obtained read dataset followed by the creation of putative unique transcripts (PUTs) with a combination of different k-mer lengths and expected coverage. Velvet program (https://www.ebi.ac.uk/~zerbino/velvet/) was used to identify the obtained transcript library (Ramadan and Hassanein, 2014). Transcripts of *C. procera*, namely *psaA*, *psaB*, *psbA* and *psbB*, were mapped to *Rhazya stricta* (Tax. ID: 396313) chloroplast complete genome (accession number KJ123753) using CLC Genomics Workbench 3.6.5. Identified sequences were blasted on GenBank (http://www.ncbi.nlm.nih.gov/BLAST) to identify available sequences with common regions.

Multi-sequence alignment (MSA) and phylogenetic analysis

To identify sequence similarities with homologous proteins from other organisms, DELTA-BLAST tools were performed to the obtained PsaA, PsaB, PsbA and PsbB deduced amino acid sequences of *C. procera*. Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) was used to generate multiple sequence alignment for the obtained transcripts. Evolutionary tree was built using the maximum-likelihood method (Saitou and Nei, 1987). CLC Genomics Workbench 8.5.1 program was used for bootstrap analysis. A bootstrap value was attached to each branch to indicate the confidence level.

The 3D homology modeling

NCBI Conserved Domain Database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) was used to identify the functional domain(s) of PsaA, Psab, PsbA and PsbB putative proteins in *C. procera*. CDD is a protein annotation resource, which uses 3D structure information to explicitly define domain boundaries and 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). Homology modeling was carried out using Protein Homology/analogY Recognition Engine V 2.0 (Phyre^2, http://www.sbg.bio.ic.ac. uk/phyre2/) program (Kelley *et al.*, 2015). Predicted model was constructed using RasMol PDB (Protein Data Bank) viewer (http://www.umass.edu/microbio/rasmol/).

Structure alignment

The protein models were applied to the pairwise comparison of protein structures using TM-align program (http://zhanglab.ccmb.med.umich.edu/TM-align/), (Zhang and Skolnick, 2005). In this program, TM-score (Zhang and Skolnick, 2004) was used to get a single scoring function that can assess the alignment quality and balance the coverage and accuracy according to the following formula:

$$TM-score = Max\left[\frac{1}{L_{Target}}\sum_{i}^{L_{ali}}\frac{1}{1+\left(\frac{d_i}{d_0(L_{Target})}\right)^2}\right]$$

where, L_{Target} is the length of target protein that other PDB structures are aligned to; L_{ali} is the number of aligned residues; d is the distance between the ith pair of aligned residues.

Quantitative real-time RT-PCR

Relative expression of the *psaA*, *psaB*, *psbA* and *psbB* transcripts of *C. procera* were determined on the RNA extracted from leaves of control and watered plants. Samples were collected from three independent but comparable desert grown plants at three time intervals; one hour after irrigation (at dawn), six hours after irrigation (at midday) and 12 hours after irrigation (one-hour pre-dusk). Control samples (Dry) were collected from the same plants one day before watering at the same time points. For each sample, 2 µg of total RNA was used to synthesize first-strand cDNA with oligo (dT) using Revert Aid Premium Reverse Transcriptase (Thermo Scientific™ cat. no. EP0451). qRT-PCR was performed with gene-specific primers (designed by GenScript Real-time PCR Primer Design, www.genscript.com/ssl-bin/app/primer). The primers used to measure the *psaA*, *psaB*, *psbA* and *psbB* gene expression are shown in Table (1). Templates were normalized to amplify 182 bp fragment of the *C. procera actin* (accession No. KU833210). The qRT-PCR was done in a total of 25-µl volume containing 1 µl cDNA, 12.5 µl 2x BIO-RAD iQTMSYBR@Green Supermix, 0.75 µl ROX reference dye (1:500 diluted), 1 µl 500 nM of each primer. All reactions were performed in triplicates. Reactions were run on Mx3005P QPCR System (Stratagene) and conditions were 1 cycle of 5 min at 95°C, 40 cycles of 30s at 95°C, 60s at 59°C, 20s at 72°C and last cycle of 72°C (infinitive). PCR products were examined by melt curve analysis. For each cDNA sample, *actin* expression levels were quantified as the reference house-keeping gene. Expression levels of each gene relative to that of *actin* gene were calculated using MxPro QPCR Software v4.10 package, which compares reaction takeoff points (cycle number). The ΔC_T for each sample was determined using the equation $\Delta C_T = C_T$ target gene – C_T reference gene to calculate the relative expression of each gene to the internal reference control (Schmittgen and Livak, 2008).

RESULTS AND DISCUSSION

The *psaA*, *psaB*, *psbA* and *psbB* transcripts of *C. procera* in the present study were recovered and mapped to *Rhazya stricta* chloroplast complete genome (accession number KJ123753) using CLC Genomics Workbench 3.6.5. The number of reads aligned for *psaA*, *psaB*, *psbA* and *psbB* transcripts were as high as 337553; 346189; 113679 and 184969, respectively. The average coverage was 13259 for *psaA* transcript, while, they were 13908, 9410 and 10631 for *psaB*, *psbA* and *psbB* transcripts, respectively. The length of consensus sequences for the four transcripts was 2253, 2205, 1062 and 1527 nt, respectively. Open reading frame (ORF) analysis showed full-length ORFs with start codons at bases 1-3, while stop codons at bases 2251-2253, 2251-2253, 1060-1062 and 1525-1527 for *psaA* (Fig. 1A), *psaB* (Fig. 1B), *psbA* (Fig. 1C) and *psbB* (Fig. 1D) transcripts, respectively. The recovered *psaA*, *psaB*, *psbA* and *psbB* transcripts (accession no. KT734792, KT734793, KT734794 and KT734795, respectively) and their deduced amino acids sequences (accession no. AML03238, AML03239, AML03240 and AML03241, respectively) were deposited in the NCBI. Identified sequences were blasted on GenBank (http://www.ncbi.nlm.nih.gov/BLAST) and related available sequences are shown in Tables (1, 2, 3 and 4, respectively).

Conserved protein domain analysis

To allocate the protein domains, the deduced amino acid sequences obtained from the ORF analysis (Fig. 1) with lengths of 751, 734, 353 and 508 aa corresponding to PsaA, PsaB, PsbA and PsbB proteins, respectively, were analyzed against the CDD database (http://www.ncbi.nlm.nih.gov/cdd). Domain analysis indicated the presence of one domain (pfam accession no. pfam00223) shared in PsaA and PsaB proteins (Fig. 2A and 2B, respectively). This domain is a member of the superfamily PsaA_PsaB domain (CDD accession no. cl08224). Structure of this domain is highly conserved in photosynthetic organisms including cyanobacteria, algae, and plants (Xiong and Bauer, 2002). PsaA and PsaB are large subunits forming the reaction center of photosystem I. They are homologous integral membrane proteins that harbor the donating chlorophyll dimer P700, chlorophyll, phylloquinone as well as the cofactors involved in light-induced electron transfer (Nelson and Ben-Shem, 2004).

Domain analysis of PsbA protein (Fig. 2C) revealed the presence of Photosystem-II_D1 domain (CDD accession no. cd09289), which is a member of Photo_RC superfamily (accession no. cl08220). PS II is a multi-subunit protein found in the photosynthetic membranes of plants, algae, and cyanobacteria (Barber, 2002). It utilizes light-induced electron transfer and water-splitting reactions to produce protons, electrons, and molecular oxygen (Goussias *et al.*, 2002). The domain contains 17 features, seven of which are binding sites and 11 are interfaces (Kawakami *et al.*, 2009). The binding sites include those of cholorophyll, pheophytin, quinone, Fe, oxygen evolving complex, bromide, beta carotene, D2 interface, CP43 interface (PSII D1 subunit interface with CP43 protein), core light harvesting interface (D1 subunit interface with the core light harvesting protein), cytochrome b559 alpha subunit interface (interface with cytochrome b559), protein I interface (PSII D1 subunit interface with reaction center protein I), protein J interface, protein L interface (PSII D1 subunit interface with reaction center protein L), manganese-stabilizing polypeptide interface (PSII D1 subunit interface with the manganese-stabilizing polypeptide), protein T interface (PSII D1 subunit interface with reaction center protein T), cytochrome c-550 interface (interface with cytochrome c-550).

Domain analysis of PsbB protein (CP47) revealed the presence of one domain belonging to PSII superfamily (CDD accession no. cl08223). The model represents a chlorophyll a antenna protein of photosystem II, which delivers light energy to photosystem II (Luciński and Jackowski, 2006).

DELTA-BLAST analysis

This analysis (http://blast.ncbi.nlm.nih.gov/) was performed to identify proteins in other organisms homologous to the PsaA, PsaB, PsbA and PsbB putative proteins in *C. procera*. The best 20 hits for the four proteins (with the lowest e-values and high identity percent) are presented in Tables (6, 7, 8 and 9, respectively). Results of the most closely-related amino acids sequences to the four *C. procera* proteins, e.g., PsaA, PsaB, PsbA and PsbB, are the Photosystem I P700 chlorophyll a apoprotein A1 of *Eustegia minuta* (e-value 0.0), photosystem I P700 chlorophyll A apoprotein A2 of *Medicago truncatula* (e-value 0.0), photosystem II thylakoid membrane protein of *Glycin emax* (e-value 4e-170) and photosystem II CP47 protein (chloroplast) of *Asclepias nivea* (e-value 0.0). These results indicate that the newly characterized PsaA and PsaB putative proteins in *C. procera* could be members of photosystem I family, while PsbA and PsbB putative proteins in *C. procera* could be members of photosystem II family.

Multi-sequence alignment (MSA) of proteins and phylogenetic analysis

MSA was performed using the best search hits for each of PsaA, PsaB, PsbA and PsbB putative proteins (Tables 6, 7, 8, and 9, respectively). Alignments for PsaA, PsaB, PsbA and PsbB putative proteins were obtained by a gap-opening penalty of 10 and a gap extension penalty of one (Figs 3, 4, 5 and 6, respectively). Data generated from MSA was used to obtain phylogenetic trees for PsaA, PsaB, PsbA and PsbB putative proteins (Fig. 7A, 7B, 7C and 7D, respectively). The results of MSA and phylogenetic tree revealed that the closest sequence to the obtained PsaA putative protein of *C. procera* is photosystem I P700 chlorophyll a apoprotein A1 of *Asclepias syriaca* (accession no. YP_008578624.1), while the closest sequence to the obtained PsaB, PsbA and PsbB putative proteins of *C. procera* are photosystem I P700 chlorophyll a apoprotein A2 of *Asclepias nivea* (accession no. YP_008578538.1), photosystem II protein D1 of *Neobracea bahamensis* (accession no. AIW05415.1) or of *Telosma cordata* (accession no. AGW04878.1) and photosystem II CP47 protein of *Matelea biflora* (accession no. AGW04691.1).

3D structural modeling

Theoretical 3D models for the studied PsaA, PsaB, PsbA and PsbB putative proteins of *C. procera* were created using the intensive mode of Phyre^2 program, (http://www.sbg.bio.ic.ac.uk/phyre^2/). This program is based on structural alignment in many stages, where a query sequence is scanned against the protein sequence database with HHblits (Kelley *et al*., 2015). Then, the recovered MSA was used to predict the secondary structure and was combined into a query hidden Markov model (HMM). The predicted structure was scanned against a database of HMMs of known structure proteins to construct crude backbone model. Then, indels in these models were corrected by loop modeling and the amino acid chains were added to generate the final model. Theoretical 3D models for PsaA, PsaB, PsbA and PsbB putative proteins of *C. procera* were created, corresponding to residues 1-750, 1-734, 1-353 and 1-508 of the primary structures, respectively (Fig. 8). The overall model dimensions are X:57.730Å/Y:71.888Å/Z:97.339Å for PsaA putative protein, while, they were X:98.590Å/Y:77.711Å/Z:72.714Å, X:68.044Å/Y:87.400Å/Z:55.703Å and X:56.873Å/Y: 71.146Å/Z: 96.334 for PsaB, PsbA and PsbB putative proteins, respectively.

Structure alignment

To determine the accuracy of the obtained 3D models, TM-align program (http://zhanglab.ccmb.med.umich.edu/TM-align/, Zhang and Skolnick, 2005) was used to compute optimal and suboptimal structural alignments between PsaA, PsaB, PsbA and PsbB 3D structures of *Arabidopsis thaliana* and *Glycine max* as compared to the theoretical 3D models of *C. procera*. The resulting superimpositions are shown in Fig. (9) and scored in Table (10). The results in Fig. (9) showed that the PsaA, PsaB, PsbA and PsbB deduced amino acids (yellow) in *C. procera* have almost the same coordination of their analogues in *A. thaliana* (blue) and *G. max* (purple), except at few positions. For PsaA protein, proline (P) residue at position 6 and aspartate residue at position 15 of *A. thaliana* and *G. max* are replaced by Alanine (A) in *C. procera*. For PsaB protein, the residues from 301-310 varied in *C. procera* versus those in *A. thaliana* and *G. max*. Also, the region 678-690 showed variations in the secondary structure, although no difference existed at the primary structure. The primary structure of PsbA was highly conserved among the three species studied; however, there were slight variations in secondary structures at residues 1-10 and 234-254 in *A. thaliana* and 1-10 and 322-327 in *G. max*. For PsbB protein, one position varied in *C. procera* versus those in *A. thaliana* (residues 488-506) and *G. max* (residues 490-508). These results support our expectation that the newly characterized PsaA and PsaB putative proteins in *C. procera* belong likely to photosystem I protein family, while PsbA and PsbB putative proteins belong likely to photosystem II protein family. The results also indicate the accuracy of the theoretical 3D models for the four studied proteins.

Differentional gene expression

To examine the effect of changes in water availability on the expression of *psaA*, *psaB*, *psbA* and *psbB* genes in *C. procera*, the transcripts abundance was determined using qRT-PCR (Fig. 10). The transcripts of *psaA* and *psaB* had their highest abundance 1 hour post dawn, then expression was decreased towards midday and pre-dusk. On the other hand, the transcripts of *psbA* and *psbB* had low abundance 1 hour post down, then expression was slightly increased at midday, reaching its maximum at pre-dusk. Comparing the transcripts abundance at each time point before and after watering, revealed the decrease in expression of the *psaA* and *psaB* transcripts in watered plants at every time point. In contrast, the *psbA* and *psbB* transcripts increased after watering, especially at pre-dusk.

Desert plants are able to survive and propagate under extremely stressful conditions, which provide a good model to study the adaptive mechanisms to a combination of stresses within their natural habitat (Mittler *et al*., 2001). Photosynthesis is affected by drought stress in terms of the decrease in CO_2 availability caused by restricted diffusion through stomata and mesophyll (Farquhar and Sharkey, 1982; Flexas *et al*., 2007). Also, non-stomatal factors had been reported, such as antenna size (Hao *et al*., 1999), in terms of the reduction of PSI and PSII electron transfer systems (Havaux *et al*., 1986).

PSI is believed to be less sensitive to abiotic stresses due to a very efficient mechanism that provide protection against photoinhibition (Tikkanen *et al*., 2014). This mechanism regulates the electron transfer chain, prevents the formation of reactive oxygen species (ROS) and protects PSI from photodamage (Tikkanen *et al*., 2014). In the present study, both *psaA* and *psaB* genes were up-regulated in plants grown under dry condition, while *psbA* and *psbB* were down-regulated under the same condition. Based on these findings, it is suggested that the balance between PSII and PSI activities affects the occurrence of (ROS). Comparable results were reached by Zhang *et al*. (2010), who found a significant increase in PSI activity and reduction of PSII reaction center in salt-stressed cyanobacterium *Spirulina platensis* as a positive response against salt stress.

PSII is also believed to play a key role under environmental stresses (Baker, 1991). Several studies demonstrated that water stress damages the reaction center of PSII (Van Rensen and Curwiel, 2000; Murata *et al*., 2007; Zlatev, 2016). Photodamage to PSII is proportional to the intensity of incident light (Takahashi and Badger, 2011). As a consequence of abiotic stress, reactive oxygen species (ROS) are generated to hamper the synthesis of the D1 protein of PSII complex leading to indirect damage of the PSII complex (Nishiyama and Murata, 2014). The decrease in PSII function (water splitting and oxygen evolving), in turn, decreases the generation of ROS (Pospísil, 2009). It was hypothesized that the decrease in PSII function maintains the integrity of the thylakoid membrane in dehydrated *Boeahygro metrica* and *Haberlear hodopensis*. Degl'Innocenti *et al*. (2008) suggested that the significant decrease in PSII activity in leaves of *Ramonda serbica* upon desiccation is a protective mechanism to maintain the membranes integrity. The core of PSII is the D1 protein, which is encoded by the chloroplast *psbA* gene. D1 protein is frequently replaced to restore the activity of damaged PS II (Nixon *et al*., 2010). The synthesis of new copies of D1 protein is a key process for the survival of plants abiotic stress-induced photoinhibition (Gururani *et al*., 2015). In the present study, the *psbA* and *psbB* genes exhibit low level of transcription rate at dry condition and at the two time sets, which exhibit a high light intensity (1 hour post-dawn and at midday). The rates of *psbA* and *psbB* gene expression were highly up-regulated in watered plants at the time of low light intensity (pre-dusk). This may enable the plant to achieve its maximum turnover rate for D1 and CP47 proteins. This could be considered as an acclimatization mechanism by which *C. procera* could achieve high rate of photosynthesis to achieve maximal growth during the rainy season. These results are in agreement with those reported by Zhang *et al*. (2002), who found a decrease in the photosynthetic gene expression in the desert plant *Ammopiptanthus mongolicus* within its natural habitat under dry condition. Upon watering, the photosynthetic genes are highly expressed and the photosynthetic assimilation rate was almost doubled.

It is evident from the aforementioned discussion on the response of *Calotropis procera* to sudden water supply that photosystem transmembrane proteins have strong impact on the mechanisms by which this desert-grown plant can stand water scarcity.

SUMMARY

The wild shrub *Calotropis procera* grows successfully in dry areas. Photosynthesis is one of the processes severely affected by drought stress. In the present study, four chloroplast genes, i.e., *psaA*, *psaB*, *psbA* and *psbB* were uncovered and characterized in *Calotropis procera* from *de novo* assembled transcriptome contigs of the high-throughput sequencing dataset. Theoretical 3D modeling of the deduced amino acid sequences was carried out and accuracy was determined by computing and suboptimal structural alignments between PsaA, PsaB, PsbA and PsbB 3D protein structures of *Arabidopsis thaliana* and *Glycine max* and the theoretical 3D models of PsaA, PsaB, PsbA and PsbB proteins in *C. procera*. Additionally, the functional domains of the studied amino acids sequences were identified. Under sudden supply of a limited amount of water to these desert grown plants, the changes in the expression of *psaA*, *psaB*, *psbA* and *psbB* genes were determined at three time points (1 hour post-dawn, midday and 1 hour pre-dusk). Data indicate that the *psaA* and *psaB* genes were down-regulated after watering, while the *psbA* and *psbB* were up-regulated especially at time point 1-hour pre-dusk. These responses can be considered as one of the mechanisms of abiotic stress tolerance in this wild plant species.

REFERENCES

Amunts, A. and N. Nelson (2009). Plant photosystem I design in the light of evolution. Structure Review, Cell Press, 17: 637-650.

Amunts, A., H. Toporik, A. Borovikova and N. Nelson (2010). Structure determination and improved model of plant photosystem I. The Journal of Biological Chemistry, 285: 3478-3486.

Amunts, A., O. Drory and N. Nelson (2007). The structure of a plant photosystem I supercomplex at 3.4 Å resolution. Nature, 447: 58-63.

Baker, N. R. (1991). A possible role for photosystem II in environmental perturbations of photosynthesis. Physiologia Plantarum, 81: 563-570.

Barber, J. (1987). Photosynthetic reaction centres : a common link. Trends Biochem. Sci., 12: 321-326.

Barber, J. (2002). Photosystem II: A multisubunit membrane protein that oxidises water. Current Opinion in Structural Biology, 12: 523-530.

Bot, A. J., F. O. Nachtergaele and A. Young (2000). Land resource potential and constraints at regional and country levels. World Soil Resources Reports, 90: 1-114.

Boutraa, T. (2010). Effects of water stress on root growth, water use efficiency, leaf area and chlorophyll content in the desert shrub *Calotropis procera*. J. Int. Environmental Application & Science, 5: 124-132.

Busch, A. and M. Hippler (2011). The structure and function of eukaryotic photosystem I. Biochimica et Biophysica Acta, 1807: 864-877.

Chaves, M. M., J. Flexas and C. Pinheiro (2009). Photosynthesis under drought and salt stress: Regulation mechanisms from whole plant to cell. Annals of Botany, 103: 551-560.

Chaves, M. M., J. P. Maroco and J. S. Pereira (2003). Understanding plant responses to drought-from genes to the whole plant. Functional Plant Biology, 30: 239-264.

Chitnis, P. R. (2001). Photosystem I: Function and physiology. Annual Review of Plant Physiology and Plant Molecular Biology, 52: 593-626.

De Souza, C. R., J. P. Maroco, T. P. dos Santos, M. L. Rodrigues, C. Lopes, J. S. Pereira and M. M. Chaves (2005). Control of stomatal aperture and carbon uptake by deficit irrigation in two grapevine cultivars. Agriculture, Ecosystems and Environment, 106: 261-274.

Degl'Innocenti, E. L. Guidi, B. Stevanovic and F. Navari (2008). CO_2 fixation and chlorophyll a fluorescence in leaves of *Ramonda serbica* during a dehydration - rehydration cycle. J. Plant Physiol., 165: 723-733.

Farquhar, G. D. and T. D. Sharkey (1982). Stomatal conductance and photosynthesis. Annu. Rev. Plant Physiol., 33: 317-345.

Flexas, J. A. Diaz-espejo, J. Galmes, R. Kaldenhoff, H. Medrano and M. Ribas-carbo (2007). Rapid variations of mesophyll conductance in response to changes in CO_2 concentration around leaves. Plant, Cell and Environment, 30: 1284-1298.

Goussias, C., A. Boussac and A. W. Rutherford (2002). Photosystem II and photosynthetic oxidation of water: an overview. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 357: 1369-1381.

Gururani, M. A., J. Venkatesh and L. P. Tran (2015). Regulation of photosynthesis during abiotic regulation of photosynthesis during abiotic stress-induced photoinhibition. Molecular Plant, 8: 1304-1320.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman and A. Regev (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols, 8: 1494-1512.

Hao, L., H. Liang, Z. Wang and X. Liu (1999). Effects of water stress and rewatering on turnover and gene expression of photosystem II reaction center polypeptide D1 in *Zea mays*. Aust. J. Plant Physiol., 26: 375-378.

Havaux, M. and K. N. Niyogi (1999). The violaxanthin cycle protects plants from photooxidative damage by more than one mechanism. Proc. Natl. Acad. Sci. USA, 96: 8762-8767.

Havaux, M., O. Canaani and S. Malkin (1986). Photosynthetic responses of leaves to water stress, expressed by photoacoustics and related methods. Plant Physiol., 82: 827-833.

Kargul, J. and J. Barber (2008). Photosynthetic acclimation: Structural reorganization of light harvesting antenna - Role of redox-dependent phosphorylation of major and minor chlorophyll a/b binding proteins. FEBS Journal, 275: 1056-1068.

Kawakami, K., Y. Umenab, N. Kamiyab and J. Shena (2009). Location of chloride and its possible functions in oxygen-evolving photosystem II revealed by X-ray crystallography. Proceedings of the National Academy of Sciences of the United States of America, 106: 8567-8572.

Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass and M. J. E. Sternberg (2015). The Phyre^2 web portal for protein modeling, prediction and analysis. Nat. Protocols, 10: 845-858.

Lawlor, D. W. and G. Cornic (2002). Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. Plant, Cell & Environment, 25: 275-294.

Loreto, F., M. Centritto and K. Chartzoulakis (2003). Photosynthetic limitations in olive cultivars with different sensitivity to salt stress. Plant, Cell and Environment, 26: 595-601.

Luciński, R. and G. Jackowski (2006). The structure, functions and degradation of pigment-binding proteins of photosystem II. Acta Biochimica Polonica, 53: 693-708.

Mittler, R. E. Merquiol, E. Hallak-Herr, S. Rachmillevitch, A. Kaplan and M. Cohen (2001). Living under a 'dormant' canopy: a molecular acclimation mechanism of the desert plant *Retama raetam*. The Plant Journal, 25: 407-416.

Mochida, K. and K. Shinozaki (2010). Genomics and bioinformatics resources for crop improvement. Plant and Cell Physiology, 51: 497-523.

Murata, N.; S. Takahashi, Y. Nishiyama and S. I. Allakhverdiev (2007). Photoinhibition of photosystem II under environmental stress. Biochim. Biophys. Acta, 1767: 414-421.

Nelson, N. and A. Ben-Shem (2004). The complex architecture of oxygenic photosynthesis. Nature Reviews, Molecular Cell Biology, 5: 1-12.

Nelson, N. and A. Ben-Shem (2005). The structure of photosystem I and evolution of photosynthesis. Bioessays, 27: 914-922.

Nishiyama, Y. and N. Murata (2014). Revised scheme for the mechanism of photoinhibition and its application to enhance the abiotic stress tolerance of the photosynthetic machinery. Appl. Microbiol. Biotechnol., 98: 8777-8796.

Nixon, P. J. F. Michoux, J. Yu, M. Boehm and J. Komenda (2010). Recent advances in understanding the assembly and repair of photosystem II. Annals of Botany, 106: 1-16.

Orwa, C., A. Mutua, R. Kindt, R. Jamnadass and A. Simons (2009). Agroforestree database: a tree reference and selection guide version 4.0, http:// www.worldagroforestry.org/af/treedb/.

Pospísil, P. (2009). Production of reactive oxygen species by photosystem II. Biochim. Biophys. Acta, 1787: 1151-1160.

Ramadan, A., J. S. M. Sabir, S. Y. M. Alakilli, A. M. Shokry, N. O. Gadalla, S. Edris, M. A. Al-Kordy, H. S. Al-Zahrani, F. M. El-Domyati, A. Bahieldin, N. R. Baker, L. Willmitzer and S. Irgang (2014). Metabolomic response of *Calotropis procera* growing in the desert to changes in water availability. PLOS ONE, 9: 13-19.

Ramadan, A. M. and S. E. Hassanein (2014). Characterization of *P5CS* gene in *Calotropis procera* plant from the de novo assembled transcriptome contigs of the high-throughput sequencing dataset. Comptes Rendus Biologies, 337: 683-690.

Rochaix, J. D. (2011). Reprint of: Regulation of photosynthetic electron transport. Biochimica et Biophysica Acta, 1807: 878-886.

Sainz, M., P. Diaz, J. Monza and O. Borsani (2010). Heat stress results in loss of chloroplast Cu/Zn superoxide dismutase and increased damage to Photosystem II in combined drought-heat stressed *Lotus japonicus*. Physiologia Plantarum, 140: 46-56.

Saitou, N. and M. Nei (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. Mol. Biol. Evol., 4: 406-425.

Schmittgen, T. D. and K. J. Livak (2008). Analyzing real-time PCR data by the comparative CT method. Nature Protocols, 3: 1101-1108.

Takahashi, S. and M. R. Badger (2011). Photoprotection in plants: a new light on photosystem II damage. Trends in Plant Science, 16: 53-60.

Tezara, W., S. Driscoll and D. W. Lawlor (2008). Partitioning of photosynthetic electron flow between CO_2 assimilation and O_2 reduction in sunflower plants under water deficit. Photosynthetica, 46: 127-134.

Tikkanen, M., N. R. Mekala and E. Aro (2014). Photosystem II photoinhibition-repair cycle protects Photosystem I from irreversible damage. Biochimica et Biophysica Acta, 1837: 210-215.

Van Rensen, J. J. and V. B. Curwiel (2000). Multiple functions of photosystem II. Indian J. Biochem. Biophys., 37: 377-382.

Xiong, J. and C. E. Bauer (2002). Complex evolution of photosynthesis. Annu. Rev. Plant. Biol., 53: 503-521.

Xu, Z., G. Zhou and H. Shimizu (2010). Plant responses to drought and rewatering. Plant Signaling & Behavior, 5: 649-654.

Zhang, L., W. Li, Y. Bi, J. Guo, J. Wen and J. Feng (2002). Water availability affects photosynthetic gene expression in the desert plant *Ammopiptanthus mongolicus*. Israel Journal of Plant Sciences, 50: 243-250.

Zhang, T., H. Gong, X. Wen and C. Lu (2010). Salt stress induces a decrease in excitation energy transfer from phycobilisomes to photosystem II but an increase to photosystem I in the cyanobacterium *Spirulina platensis*. Journal of Plant Physiology, 167: 951-958.

Zhang, Y. and J. Skolnick (2004). Scoring function for automated assessment of protein structure template quality. Proteins, Structure, Function and Genetics, 57: 702-710.

Zhang, Y. and J. Skolnick (2005). TM-align: A protein structure alignment algorithm based on the TM-score. Nucleic Acids Research, 33: 2302-2309.

Zlatev, Z. (2016). Drought-induced changes in chlorophyll fluorescence of young wheat plants. Biotechnol. Biotechnol. Equip., 23: 438-441.

Table (1): Primer names and sequences to amplify *psaA*, *psaB*, *psbA* and *psbB* transcript sequences along with those to amplify *actin* gene sequence used as the house-keeping gene.

No.	Primer name	Primer sequence
1	*psaA* (Forward)	5`TCGGCTTGGGATCATGTCTT 3`
	psaA (Reverse)	5`AGAAATCACGTAGCCACCCA 3`
2	*psaB* (Forward)	5`ATAGTTTGTCGGTCTGGGCA 3`
	psaB (Reverse)	5`TGCTTGCACAATGGAAAGGG 3`
3	*psbA* (Forward)	5`TTACATGGGTCGTGAGTGGG 3`
	psbA (Reverse)	5`GTTGTGCTCAGCCTGGAATAC 3`
5	*psbB* (Forward)	5`CGGCCGGTTACTTTCTGTTC 3`
	psbB (Reverse)	5`ACTCCAACCGCTCCATGAAT 3`
6	*actin* (Forward)	5`GCACACTGGTGTCATGGTTG 3`
	Actin (Reverse)	5`CCTCAGGAGCAACACGAAGT 3`

Table (2): Accession numbers and description of the genes and organisms with highest similarities to *psaA* transcript (accession no. KT734792) of *C. procera*.

Accession no.	Description/Organism	Score	E-value
AGW04975.1	photosystem I P700 chlorophyll a apoprotein A1 [*Vincetoxicum rossicum*]	1452	0.0
AGW04667.1	photosystem I P700 chlorophyll a apoprotein A1 [*Matelea biflora*]	1452	0.0
AGW04441.1	photosystem I P700 chlorophyll a apoprotein A1 [*Astephanus triflorus*]	1451	0.0
YP_008578624.1	photosystem I P700 chlorophyll a apoprotein A1 [*Asclepias syriaca*]	1450	0.0
AGW04590.1	photosystem I P700 chlorophyll a apoprotein A1 [*Marsdenia astephanoides*]	1450	0.0
AGW04518.1	photosystem I P700 chlorophyll a apoprotein A1 [*Eustegia minuta*]	1449	0.0
AGW04821.1	photosystem I P700 chlorophyll a apoprotein A1 [*Sisyranthus trichostomus*]	1447	0.0
AGW04744.1	photosystem I P700 chlorophyll a apoprotein A1 [*Orthosia scoparia*]	1447	0.0
AIW05774.1	photosystem I P700 chlorophyll a apoprotein A1 [*Periploca sepium*]	1444	0.0
YP_009108843.1	photosystem I P700 chlorophyll a apoprotein A1 [*Pentalinon luteum*]	1444	0.0
YP_008081265.1	photosystem I P700 apoprotein A1 (chloroplast) [*Catharanthus roseus*]	1444	0.0
YP_008592488.1	photosystem I P700 chlorophyll a apoprotein A1 [*Andrographis paniculata*]	1441	0.0
AJE71806.1	photosystem I P700 chlorophyll a apoprotein A1 [*Amorpha canescens*]	1440	0.0
AKC98674.1	photosystem I P700 apoprotein A1 [*Corymbia citriodora* subsp.*variegate*]	1439	0.0
YP_009117221.1	photosystem I P700 apoprotein A1 [*Premna microphylla*]	1439	0.0
NP_051059.1	photosystem I P700 chlorophyll a apoprotein A1 [*Arabidopsis thaliana*]	1436	0.0
YP_009123073.1	photosystem I P700 apoprotein A1 [*Cannabis sativa*]	1436	0.0
YP_005089952.1	psaA gene product [*Brassica napus*]	1434	0.0
YP_538935.1	photosystem I P700 apoprotein A1 [*Gossypium hirsutum*]	1432	0.0
YP_538755.1	photosystem I P700 apoprotein A1 [*Glycine max*]	1422	0.0

Table (3): Accession numbers and description of the genes and organisms with highest similarities to *psaA* transcript (accession no. KT734793) of *C. procera*.

Accession no.	Description / Organism	Score	E-value
YP_008578538.1	photosystem I P700 chlorophyll a apoprotein A2 (chloroplast) [*Asclepias nivea*]	1434	0.0
AGW04974.1	photosystem I P700 chlorophyll a apoprotein A2 [*Vincetoxicum rossicum*]	1432	0.0
YP_009162260.1	photosystem I P700 apoprotein A2 [*Scutellaria baicalensis*]	1432	0.0
YP_008081264.1	photosystem I P700 apoprotein A2 (chloroplast) [*Catharanthus roseus*]	1431	0.0
YP_003359358.1	PSI P700 apoprotein A2 [*Olea europaea*]	1430	0.0
NP_054496.1	photosystem I P700 chlorophyll a apoprotein A2 [*Nicotiana tabacum*]	1429	0.0
AKZ22663.1	photosystem I P700 chlorophyll a apoprotein A2 [*Solanum rostratum*]	1429	0.0
YP_567075.1	photosystem I P700 apoprotein A2 [*Vitis vinifera*]	1427	0.0
YP_009171867.1	PsaB [*Solanum nigrum*]	1426	0.0
AJP62109.1	photosystem I P700 chlorophyll a apoprotein A2 [*Dianthus longicalyx*]	1424	0.0
AGW98119.1	photosystem I P700 apoprotein A2 [*Ipomoea ternifolia*]	1423	0.0
YP_004940509.1	psaB gene product [*Boea hygrometrica*]	1422	0.0
YP_009185342.1	photosystem I P700 apoprotein A2 [*Tilia amurensis*]	1419	0.0
YP_009132938.1	photosystem I P700 apoprotein A2 [*Hibiscus syriacus*]	1419	0.0
YP_913186.1	PSI P700 apoprotein A2 [*Gossypium barbadense*]	1418	0.0
AJE73243.1	photosystem I P700 chlorophyll a [*Bidens aristosa*]	1417	0.0
NP_051058.1	photosystem I P700 chlorophyll a apoprotein A2 [*Arabidopsis thaliana*]	1414	0.0
YP_009161022.1	PsaB [*Lilium hansonii*]	1413	0.0
XP_003599572.2	photosystem I P700 chlorophyll A apoprotein A2 [*Medicago truncatula*]	1409	0.0
YP_538756.1	photosystem I P700 apoprotein A2 [*Glycine max*]	1407	0.0

Table (4): Accession numbers and description of the genes and organisms with highest similarities to *psbA* transscript (accession No. KT734794) of *C. procera*.

Accession no.	Description /Organism	Score	E-value
YP_008578519.1	photosystem II protein D1 (chloroplast) [*Asclepias nivea*]	685	0.0
YP_005089932.1	psbA gene product [*Brassica napus*]	684	0.0
YP_009170050.1	photosystem II protein D1 [*Larrea tridentata*]	684	0.0
AER52639.1	photosystem II protein D1 [*Asclepias cutleri*]	684	0.0
YP_004072442.1	PSII 32 kDa protein [*Corynocarpus laevigata*]	684	0.0
YP_009160782.1	photosystem II protein D1 [*Haloxylon ammodendron*]	684	0.0
YP_538745.1	photosystem II protein D1 [*Glycine max*]	684	0.0
AIW05415.1	photosystem II protein D1 [*Neobracea bahamensis*]	684	0.0
YP_009162857.1	photosystem II protein D1 [*Rheum palmatum*]	684	0.0
YP_003934346.1	photosystem II protein D1 [*Monsonia speciosa*]	684	0.0
YP_009046894.1	photosystem II protein D1 [*Raphanus sativus*]	684	0.0
YP_008081245.1	photosystem II protein D1 (chloroplast) [*Catharanthus roseus*]	684	0.0
YP_001122788.2	photosystem II protein D1 [*Phaseolus vulgaris*]	684	0.0
NP_051039.1	photosystem II protein D1 [*Arabidopsis thaliana*]	684	0.0
ABH88068.1	photosystem II protein D1 [*Phaseolus vulgaris*]	683	0.0
AAQ67339.1	photosystem II thylakoid membrane protein [*Glycine max*]	683	0.0
YP_003934108.1	photosystem II protein D1 [*Geranium palmatum*]	683	0.0
AGW04878.1	photosystem II protein D1 [*Telosma cordata*]	683	0.0
YP_006503771.1	photosystem II protein D1 [*Datura stramonium*]	682	0.0

Table (5): Accession numbers and description of the genes and organisms with highest similarities to *psbB* transcript (accession no. KT734795) of *C. procera*.

Accession no.	Description / Organism	Score	E-value
AGW04999.1	photosystem II CP47 protein [*Vincetoxicum rossicum*]	1000	0.0
YP_008578565.1	photosystem II CP47 protein (chloroplast) [*Asclepias nivea*]	1000	0.0
AGW04691.1	photosystem II CP47 protein [*Matelea biflora*]	999	0.0
AGW04614.1	photosystem II CP47 protein [*Marsdenia astephanoides*]	999	0.0
AGW04388.1	photosystem II CP47 protein [*Araujia sericifera*]	999	0.0
AER53249.1|	photosystem II CP47 protein [*Asclepias subaphylla*]	999	0.0
AER52847.1	photosystem II CP47 protein [*Asclepias leptopus*]	999	0.0
AER52684.1	photosystem II CP47 protein [*Asclepias cutleri*]	999	0.0
AGW04465.1	photosystem II CP47 protein [*Astephanus triflorus*]	996	0.0
AGW04845.1	photosystem II CP47 protein [*Sisyranthus trichostomus*]	994	0.0
AGW04541.1	photosystem II CP47 protein [*Eustegia minuta*]	994	0.0
YP_009183618.1	PSII 47 kDa protein [*Scutellaria insignis*]	985	0.0
YP_009144540.1	photosystem II 47 kDa protein [*Rosmarinus officinalis*]	984	0.0
YP_004935692.1	photosystem II 47 kDa protein [*Sesamum indicum*]	984	0.0
NP_054526.1	photosystem II 47 kDa protein [*Nicotiana tabacum*]	981	0.0
YP_009132965.1	photosystem II 47 kDa protein [*Hibiscus syriacus*]	981	0.0
ALI90542.1	PsbB [*Pittosporopsis kerrii*]	981	0.0
NP_051084.1	photosystem II 47 kDa protein [*Arabidopsis thaliana*]	978	0.0
YP_538791.1	photosystem II 47 kDa protein [*Glycine max*]	967	0.0

Table (6): Accession numbers, description and organisms and the calculated E-value of homologous proteins to PsaA deduced amino acids sequence (accession no. AML03238) of *C. procera* identified using DELTA-BLAST search program.

Accession no.	Description / Organism	Total score	Query cover (%)	E-value	Max. ident. (%)
AGW04518.1	photosystem I P700 chlorophyll a apoprotein A1[*Eustegia minuta*]	1192	99	0.0	99
YP_008578624.1	photosystem I P700 chlorophyll a apoprotein A1 [*Asclepias syriaca*]	1191	99	0.0	99
AGW04821.1	photosystem I P700 chlorophyll a apoprotein A1 [*Sisyranthus trichostomus*]	1191	99	0.0	99
AJE71806.1	photosystem I P700 chlorophyll a apoprotein A1 [*Amorpha canescens*]	1190	99	0.0	98
YP_009108843.1	photosystem I P700 chlorophyll a apoprotein A1 [*Pentalinon luteum*]	1190	99	0.0	99
AIW05774.1	photosystem I P700 chlorophyll a apoprotein A1 [*Periploca sepium*]	1190	99	0.0	99
AGW04975.1	photosystem I P700 chlorophyll a apoprotein A1 [*Vincetoxicum rossicum*]	1190	99	0.0	99
AGW04667.1	photosystem I P700 chlorophyll a apoprotein A1 [*Matelea biflora*]	1190	99	0.0	99
AGW04590.1	photosystem I P700 chlorophyll a apoprotein A1[*Marsdenia astephanoides*]	1190	99	0.0	99
AGW04441.1	photosystem I P700 chlorophyll a apoprotein A1[*Astephanus triflorus*]	1190	99	0.0	99
YP_008592488.1	photosystem I P700 chlorophyll a apoprotein A1[*Andrographis paniculata*]	1190	99	0.0	98
YP_538935.1	photosystem I P700 apoprotein A1[*Gossypium hirsutum*]	1190	99	0.0	98
AKC98674.1	photosystem I P700 apoprotein A1[*Corymbia citriodora subsp. Variegate*]	1189	99	0.0	98
YP_009117221.1	photosystem I P700 apoprotein A1[*Premna microphylla*]	1189	99	0.0	98
AGW04744.1	photosystem I P700 chlorophyll a apoprotein A1[*Orthosia scoparia*]	1189	99	0.0	99
YP_005089952.1	*psaA* gene product [*Brassica napus*]	1189	99	0.0	99
YP_009123073.1	photosystem I P700 apoprotein A1[*Cannabis sativa*]	1189	99	0.0	98
NP_051059.1	photosystem I P700 chlorophyll a apoprotein A1[*Arabidopsis thaliana*]	1187	99	0.0	98
YP_008081265.1	photosystem I P700 apoprotein A1 [*Catharanthus roseus*]	1186	99	0.0	99
YP_538755.1	photosystem I P700 apoprotein A1[*Glycine max*]	1183	99	0.0	97

Table (7): Accession numbers, description and organisms and the calculated E-value of homologous proteins to PsaB deduced amino acids sequence (accession no. AML03239) of *C. procera* identified using DELTA-BLAST search program.

Accession no.	Description /Organism	Total score	Query cover (%)	E-value	Max ident %
XP_003599572.2	photosystem I P700 chlorophyll A apoprotein A2 [*Medicago truncatula*]	1130	100	0.0	96
YP_008578538.1	photosystem I P700 chlorophyll a apoprotein A2 [*Asclepias nivea*]	1124	100	0.0	99
YP_004940509.1	psaB gene product [*Boea hygrometrica*]	1124	100	0.0	98
YP_009185342.1	photosystem I P700 apoprotein A2 [*Tilia amurensis*]	1122	100	0.0	98
YP_009132938.1	photosystem I P700 apoprotein A2 [*Hibiscus syriacus*]	1122	100	0.0	98
AGW98119.1	photosystem I P700 apoprotein A2 [*Ipomoea ternifolia*]	1122	100	0.0	98
AGW04974.1	photosystem I P700 chlorophyll a apoprotein A2 [*Vincetoxicum rossicum*]	1122	100	0.0	99
YP_008081264.1	photosystem I P700 apoprotein A2 [*Catharanthus roseus*]	1122	100	0.0	99
YP_913186.1	PSI P700 apoprotein A2 [*Gossypium barbadense*]	1122	100	0.0	98
AJE73243.1	photosystem I P700 chlorophyll a [*Bidens aristosa*]	1122	100	0.0	98
NP_054496.1	photosystem I P700 chlorophyll a apoprotein A2 [*Nicotiana tabacum*]	1122	100	0.0	99
YP_009162260.1	photosystem I P700 apoprotein A2 [*Scutellaria baicalensis*]	1122	100	0.0	99
AJP62109.1	photosystem I P700 chlorophyll a apoprotein A2 [*Dianthus longicalyx*]	1122	100	0.0	98
YP_003359358.1	PSI P700 apoprotein A2 [*Olea europaea*]	1122	100	0.0	99
YP_567075.1	photosystem I P700 apoprotein A2 [*Vitis vinifera*]	1122	100	0.0	99
AKZ22663.1	photosystem I P700 chlorophyll a apoprotein A2 [*Solanum rostratum*]	1122	100	0.0	99
YP_009161022.1	PsaB [*Lilium hansonii*]	1121	100	0.0	97
YP_009171867.1	PsaB [*Solanum nigrum*]	1120	100	0.0	99
YP_538756.1	photosystem I P700 apoprotein A2 [*Glycine max*]	1114	100	0.0	97
NP_051058.1	photosystem I P700 chlorophyll a apoprotein A2 [*Arabidopsis thaliana*]	1107	100	0.0	97

Table (8): Accession numbers, description and organisms and the calculated E-value of homologous proteins to PsbA deduced amino acids sequence (accession No. AML03240) of *C. procera* identified using DELTA-BLAST search program.

Accession no.	Description /Organism	Total score	Query cover (%)	E-value	Max ident %
AAQ67339.1	photosystem II thylakoid membrane protein [*Glycine max*]	491	100	4e-170	99
ABH88068.1	photosystem II protein D1 [*Phaseolus vulgaris*]	490	100	5e-170	99
YP_008578519.1	photosystem II protein D1[*Asclepias nivea*]	490	100	6e-170	100
YP_005089932.1	psbA gene product [*Brassica napus*]	490	100	7e-170	99
BAG70962.1	D1 protein [*Ambrosia artemisiifolia*]	489	100	8e-170	99
YP_003934346.1	photosystem II protein D1 [*Monsonia speciosa*]	489	100	9e-170	99
AGW04878.1	photosystem II protein D1 [*Telosma cordata*]	489	100	1e-169	99
AER52639.1	photosystem II protein D1 [*Asclepias cutleri*]	489	100	1e-169	99
YP_538745.1	photosystem II protein D1 [*Glycine max*]	489	100	1e-169	99
YP_009046894.1	photosystem II protein D1 [*Raphanus sativus*]	489	100	1e-169	99
YP_009160782.1	photosystem II protein D1 [*Haloxylon ammodendron*]	489	100	1e-169	99
AIW05415.1	photosystem II protein D1 [*Neobracea bahamensis*]	489	100	2e-169	99
YP_009162857.1	photosystem II protein D1 [*Rheum palmatum*]	489	100	2e-169	99
YP_008081245.1	photosystem II protein D1 [*Catharanthus roseus*]	489	100	2e-169	99
YP_004072442.1	PSII 32 kDa protein [*Corynocarpus laevigata*]	489	100	2e-169	99
YP_003934108.1	photosystem II protein D1 [*Geranium palmatum*]	489	100	2e-169	99
YP_009170050.1	photosystem II protein D1 [*Larrea tridentata*]	489	100	2e-169	99
YP_001122788.2	photosystem II protein D1 [*Phaseolus vulgaris*]	489	100	2e-169	99
YP_006503771.1	photosystem II protein D1 [*Datura stramonium*]	489	100	2e-169	99
NP_051039.1	photosystem II protein D1 [*Arabidopsis thaliana*]	488	100	2e-169	99

Table (9): Accession numbers, description and organisms and the calculated E-value of homologous proteins to PsbB deduced amino acids sequence (accession No. AML03241) of *C. procera* identified using DELTA-BLAST search program.

Accession no.	Description /Organism	Total score	Query cover (%)	E-value	Max ident %
YP_008578565.1	photosystem II CP47 protein (chloroplast) [*Asclepias nivea*]	791	100	0.0	99
AGW04999.1	photosystem II CP47 protein [*Vincetoxicum rossicum*]	791	100	0.0	99
AER52847.1	photosystem II CP47 protein [*Asclepias leptopus*]	790	100	0.0	99
AER53249.1	photosystem II CP47 protein [*Asclepias subaphylla*]	790	100	0.0	99
AER52684.1	photosystem II CP47 protein [*Asclepias cutleri*]	790	100	0.0	99
AGW04691.1	photosystem II CP47 protein [*Matelea biflora*]	789	100	0.0	99
AGW04614.1	photosystem II CP47 protein [*Marsdenia astephanoides*]	788	100	0.0	99
AGW04388.1	photosystem II CP47 protein [*Araujia sericifera*]	788	100	0.0	99
AGW04845.1	photosystem II CP47 protein [*Sisyranthus trichostomus*]	787	100	0.0	99
AGW04465.1	photosystem II CP47 protein [*Astephanus triflorus*]	787	100	0.0	99
ALI90542.1	PsbB [*Pittosporopsis kerrii*]	786	100	0.0	99
AGW04541.1	photosystem II CP47 protein [*Eustegia minuta*]	785	100	0.0	99
YP_008578114.1	photosystem II 47 kDa protein [*Allosyncarpia ternata*]	784	100	0.0	97
YP_004935692.1	photosystem II 47 kDa protein [*Sesamum indicum*]	784	100	0.0	98
YP_009046940.1	photosystem II 47 kDa protein [*Raphanus sativus*]	783	100	0.0	96
YP_009183618.1	PSII 47 kDa protein [*Scutellaria insignis*]	781	100	0.0	98
YP_002720138.1	psbB [*Jatropha curcas*]	778	100	0.0	97
NP_054526.1	photosystem II 47 kDa protein [*Nicotiana tabacum*]	778	00	0.0	97
NP_051084.1	photosystem II 47 kDa protein [*Arabidopsis thaliana*]	777	100	0.0	96
YP_538791.1	photosystem II 47 kDa protein [*Glycine max*]	767	100	0.0	95

Table (10): Aligned residues (AR), TM-align score (TM-score) and root-mean-square deviation (RMSD) of pairwise structural alignment of PsaA, PsaB, PsbA and PsbB proteins in *C. procera* and those of *Arabidopsis thaliana* and *Glycine max*.

	Arabidopsis thaliana			*Glycine max*		
	AR	TM-score	RMSD	AR	TM-score	RMSD
PsaA	784	0.99220	0.77	749	0.99458	0.78
PsaB	734	0.98653	1.23	734	0.98772	1.12
PsbA	344	0.91689	2.07	346	0.94206	1.84
PsbB	493	0.95620	1.23	501	0.97189	1.26

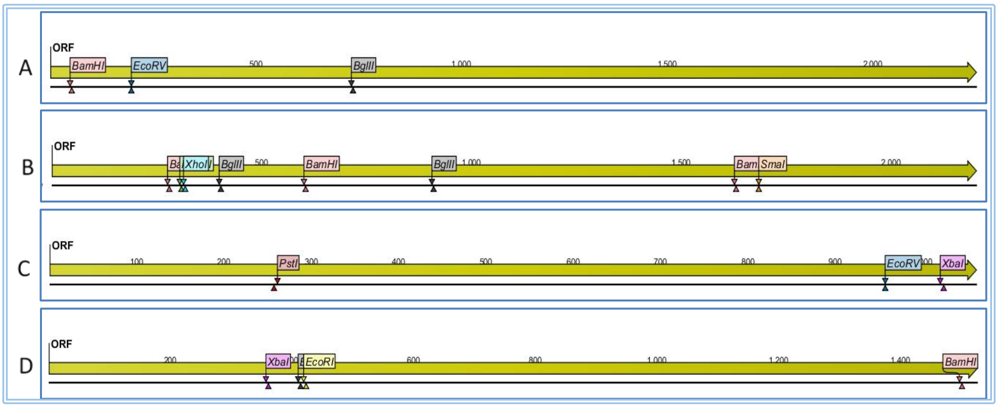

Fig. (1): Open reading frame (ORF) analysis indicating the restriction maps for the obtained *psaA*, *psaB*, *psbA*, *psbB* photosystem I and II genes in *C. procera*. A; *psaA* gene sequence is characterized by the presence of *BamHI*, *EcoRV* and *BglII* sites at 48, 197 and 732 nt, respectively. B; *psaB* gene sequence is characterized by the presence of two *BamH1* sites at 276 and 601 nt; *HindIII* at 305 nt, *XhoI* at 314 nt and *BglII* sites at 399 and 906 nt. C; *psbA* gene sequence is characterized by the presence of *XbaI*, *PstI* and *EcoRV* sites at 120, 261 and 957 nt, respectively. D; *psbB* gene sequence is characterized by the presence of *XbaI*, *BglII* and *EcoR1* sites at 357, 410 and 419 nt, respectively.

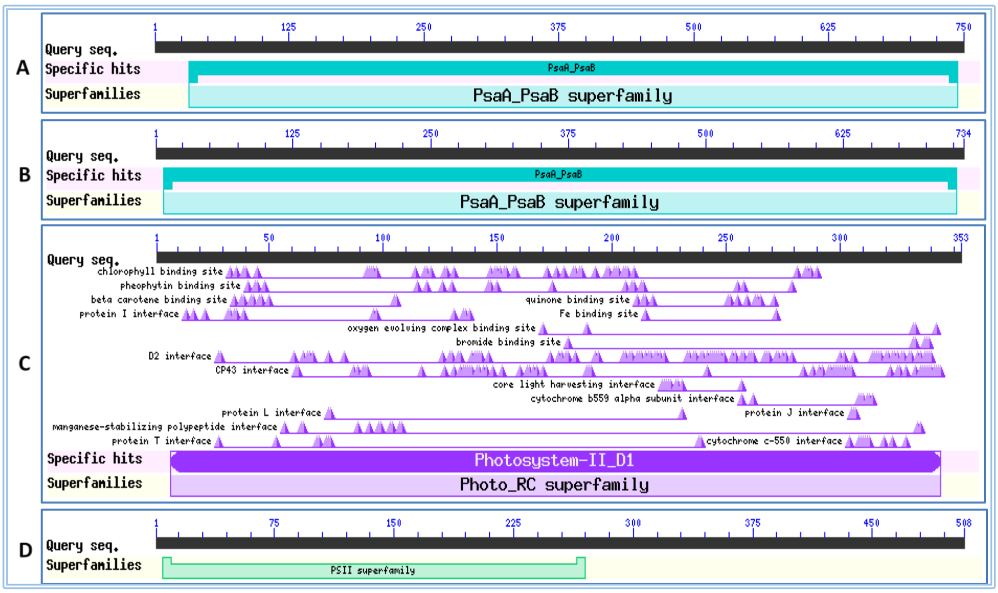

Fig. (2): Protein domains for the obtained four deduced amino acid sequences of the *psaA* (A), *psaB* (B), *psbA* (C) and *psbB* (D) photosystem I and II genes in *C. procera* as analyzed by pfam database.

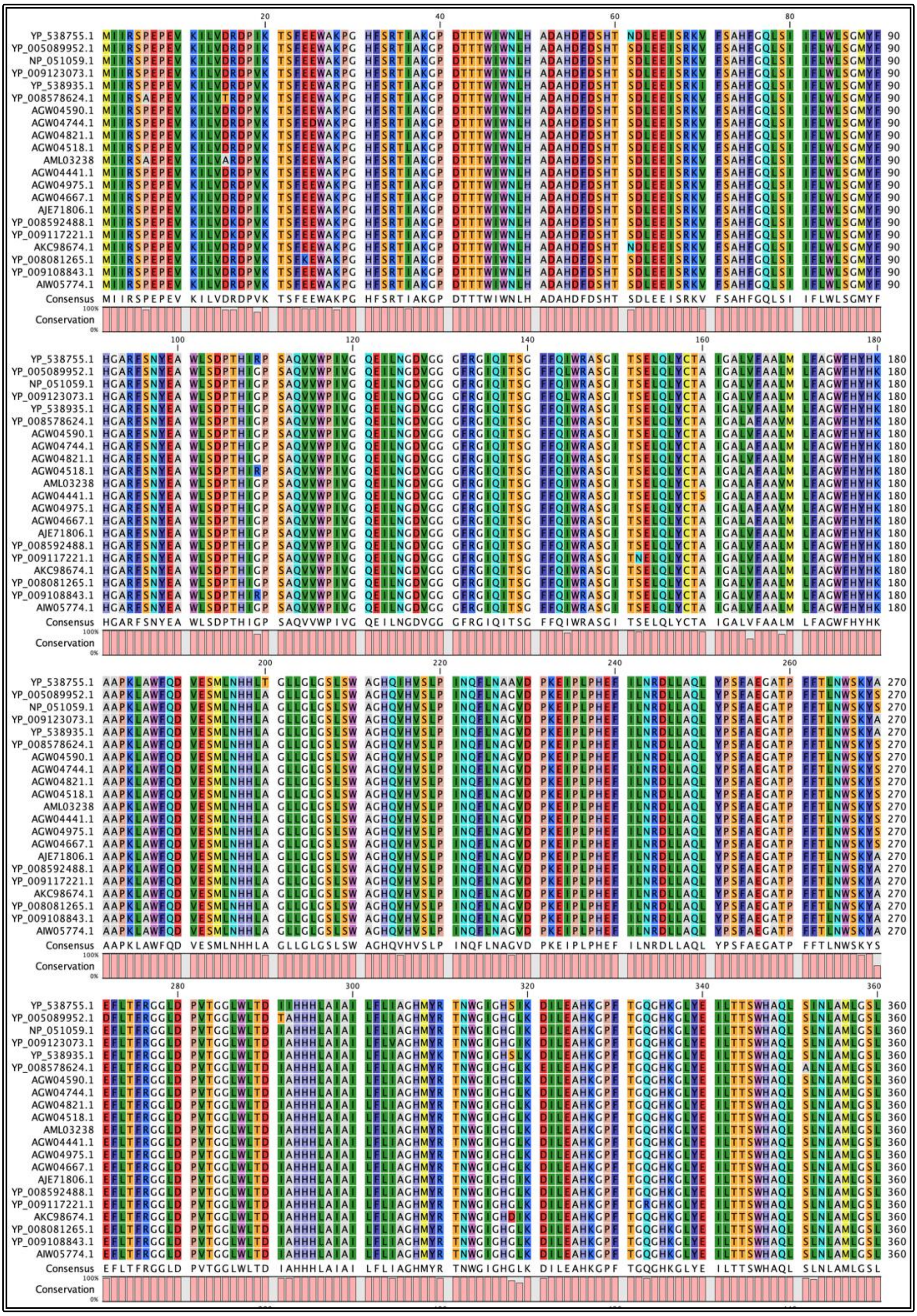

Fig. (3): Multiple sequence alignment of the 20 different PsaA protein sequences with the obtained PsaA deduced amino acids sequence of *C. procera* (accession No. AML03238).

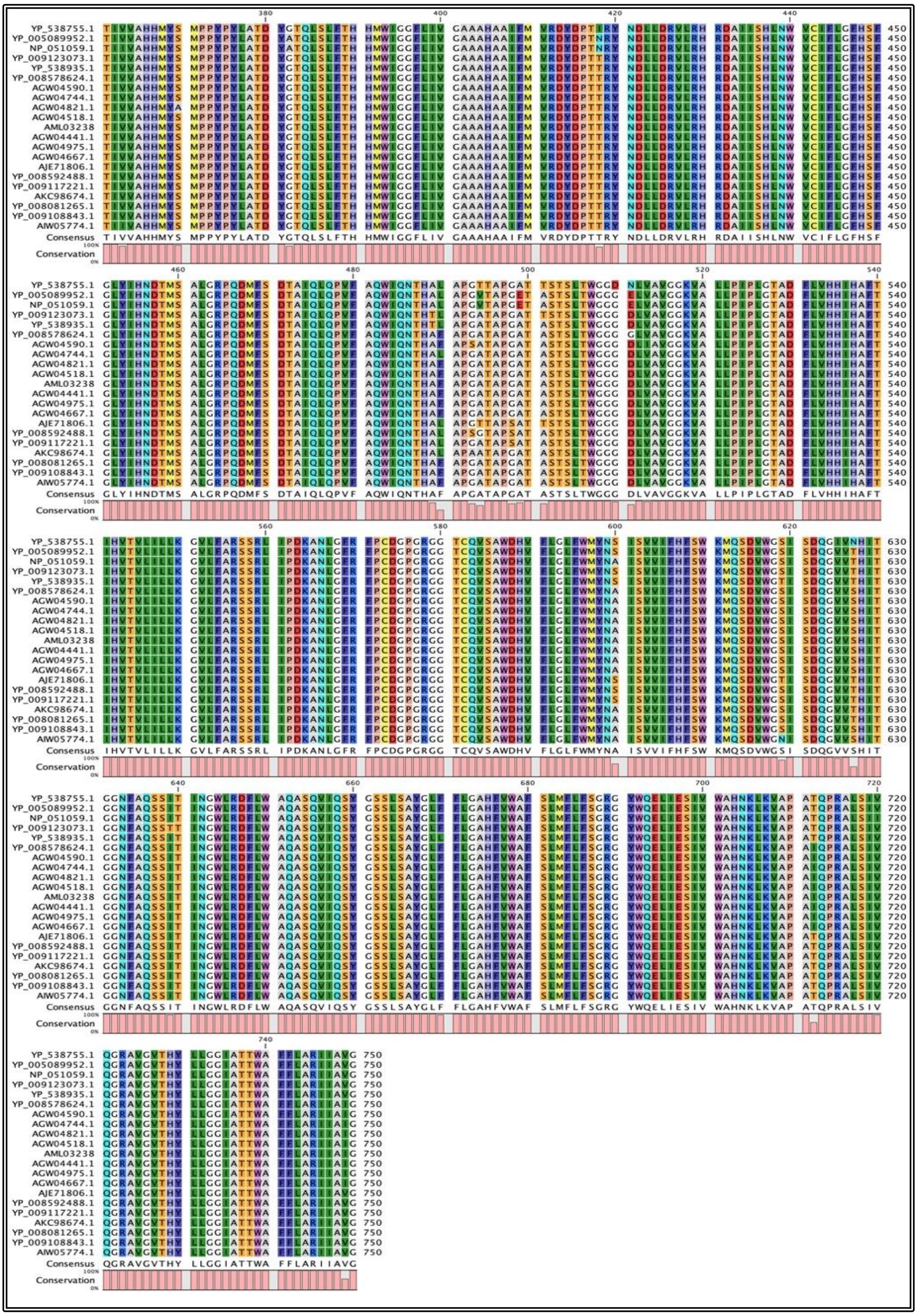

Fig. (3): Cont.

Fig. (4): Multiple sequence alignment of the 20 different PsaB protein sequences with the obtained PsaB deduced amino acids sequence of *C. procera* (accession No. AML03239).

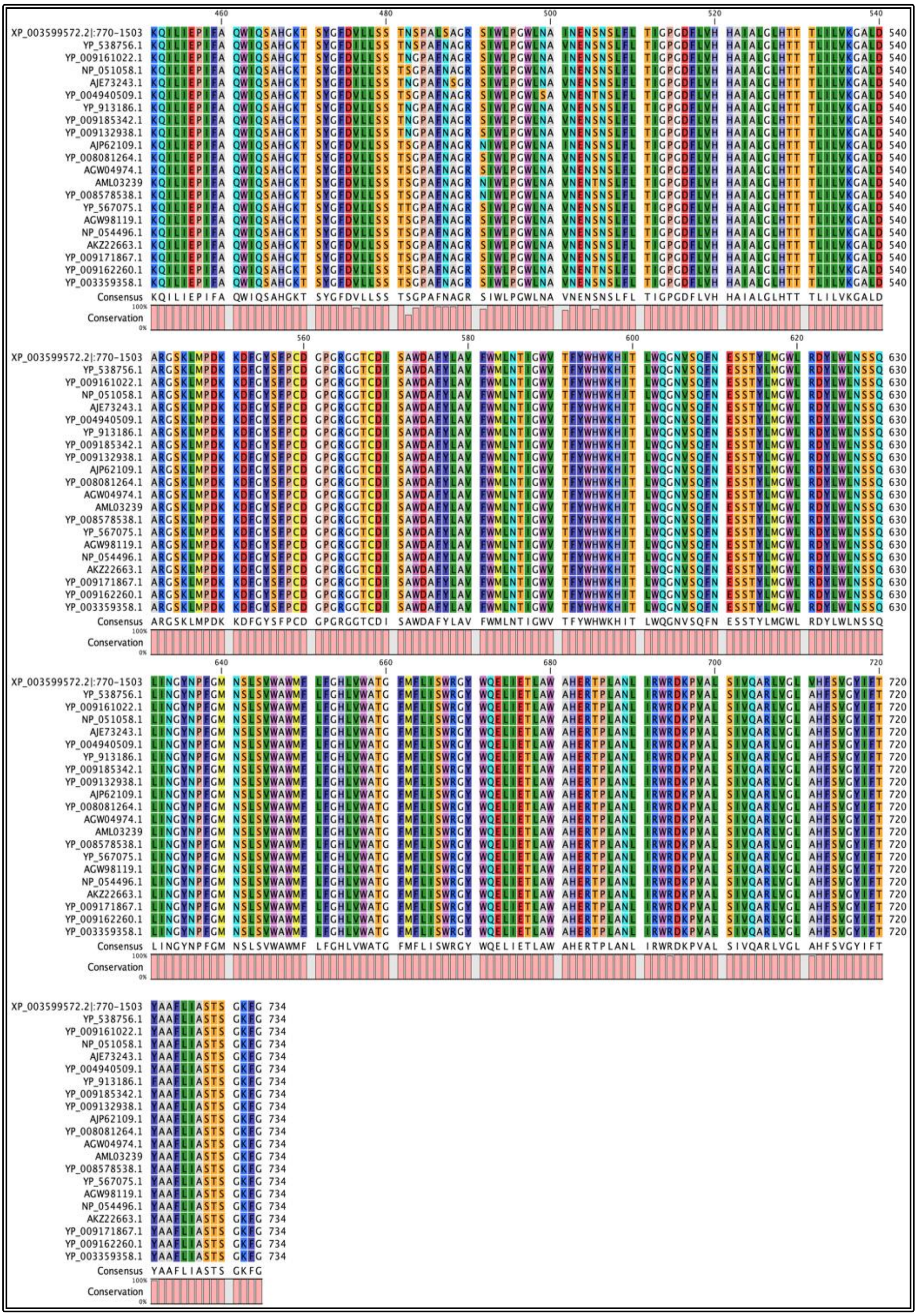
Consensus KQILIEPIFA QWIQSAHGKT SYGFDVLLSS TSGPAFNAGR SIWLPGWLNA VNENSNSLFL TIGPGDFLVH HAIALGLHTT TLILVKGALD
Conservation
Consensus ARGSKLMPDK KDFGYSFPCD GPGRGGTCDI SAWDAFYLAV FWMLNTIGWV TFYWHWKHIT LWQGNVSQFN ESSTYLMGWL RDYLWLNSSQ
Conservation
Consensus LINGYNPFGM NSLSVWAWMF LFGHLVWATG FMFLISWRGY WQELIETLAW AHERTPLANL IRWRDKPVAL SIVQARLVGL AHFSVGYIFT
Conservation
Consensus YAAFLIASTS GKFG
Conservation

Fig. (4): Cont.

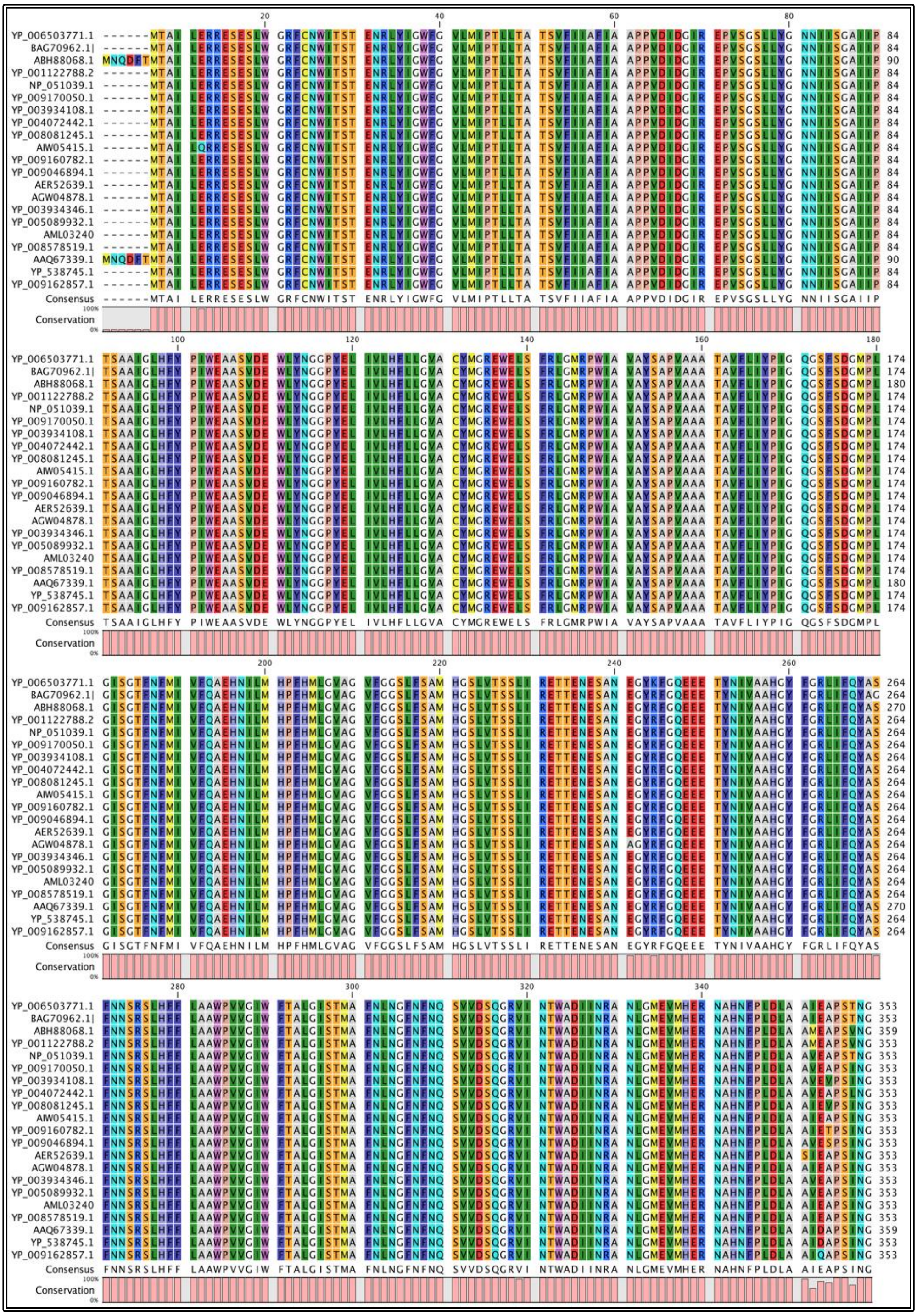

Fig. (5): Multiple sequence alignment of the 20 different PsbA protein sequences with the obtained PsbA deduced amino acids sequence of *C. procera* (accession No. AML03240).

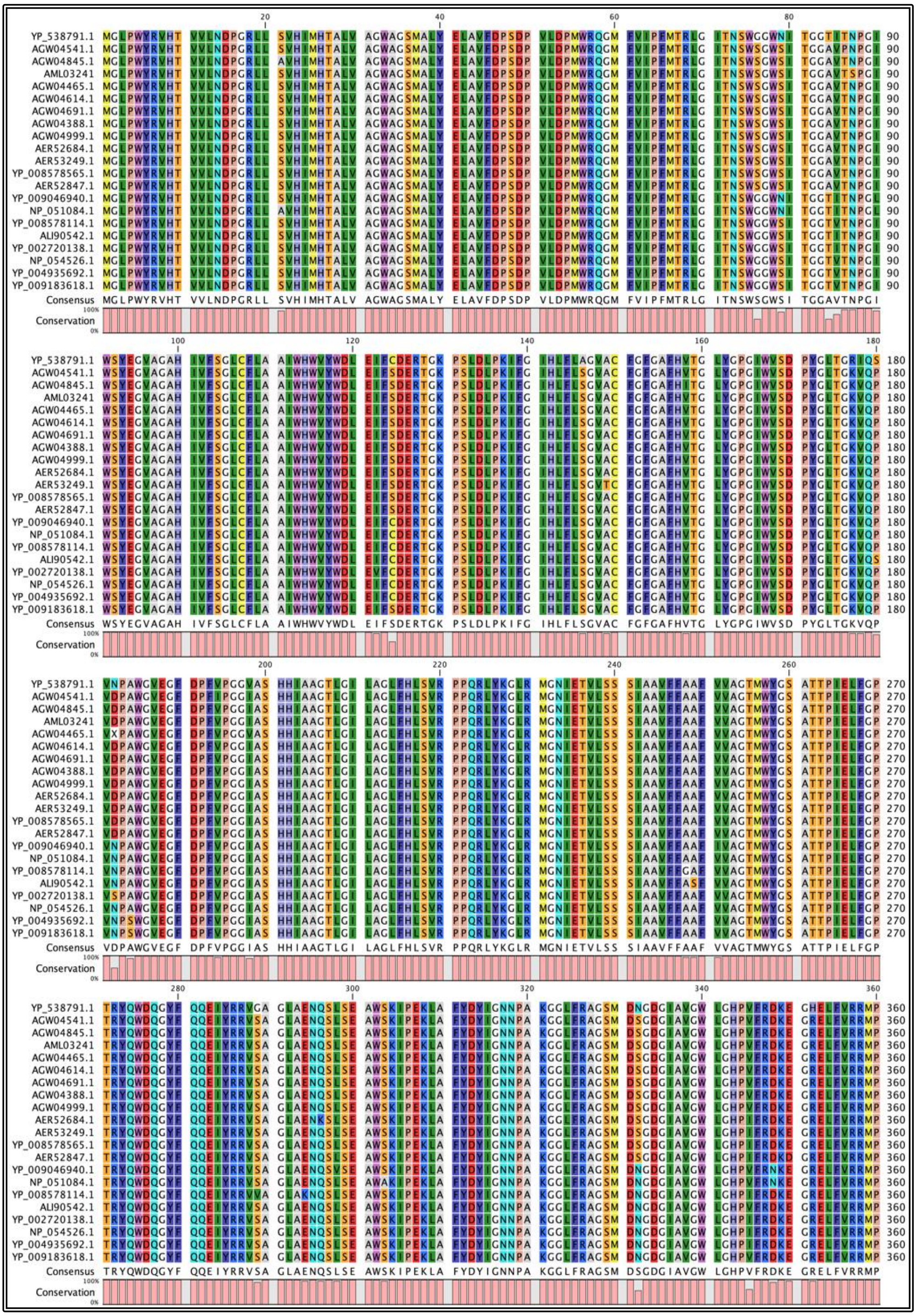

Fig. (6): Multiple sequence alignment of the 20 different PsbB protein sequences with the obtained PsbB deduced amino acids sequence of *C. procera* (accession No. AML03241).

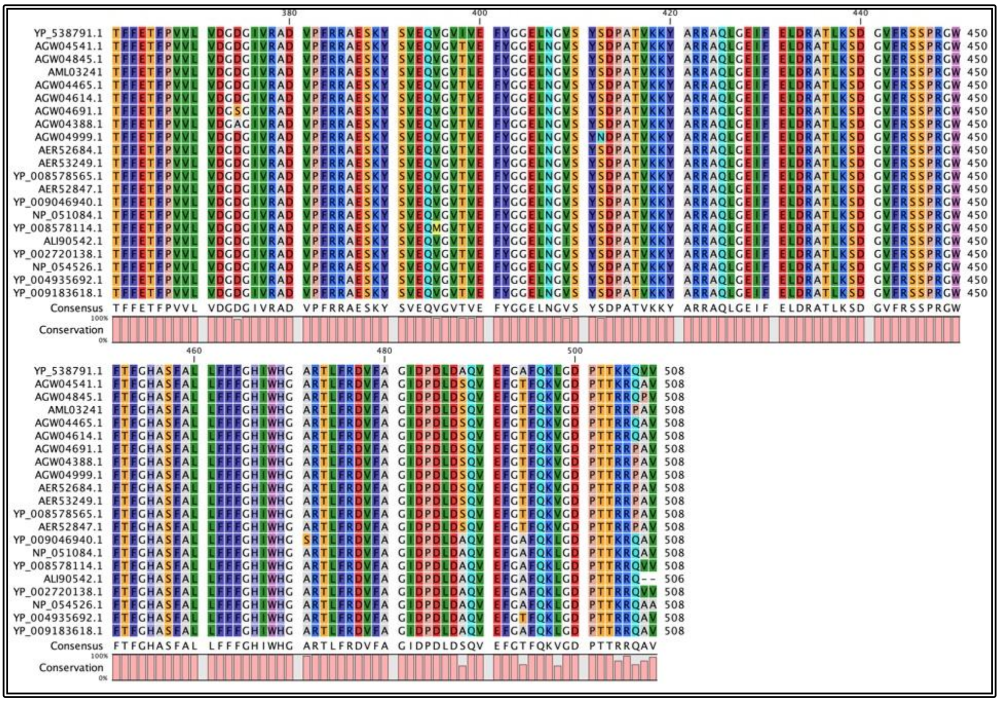

Fig. (6): Cont.

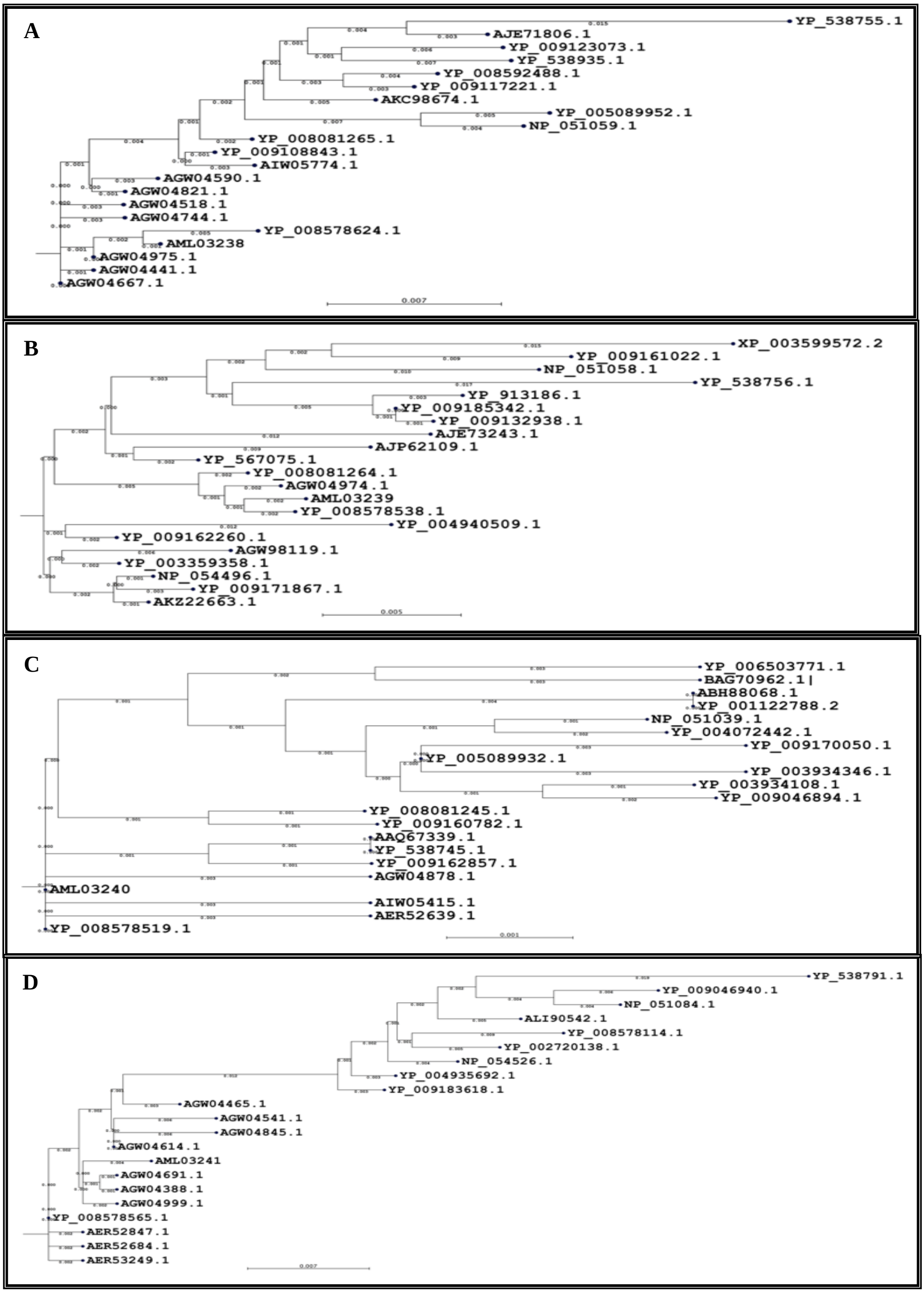

Fig. (7): Phylogenetic analysis of: A; 20 PsaA proteins and PsaA deduced amino acids sequence (accession no. AML03238) of *C. procera*. B; 20 PsaB proteins and PsaB deduced amino acids sequence (accession no. AML03239) of *C. procera*. C; 20 PsbA proteins and PsbA deduced amino acids sequence (accession no. AML03240) of *C. procera*. D; 20 PsbB proteins and PsbB deduced amino acids sequence (accession no. AML03241) of *C. procera*.

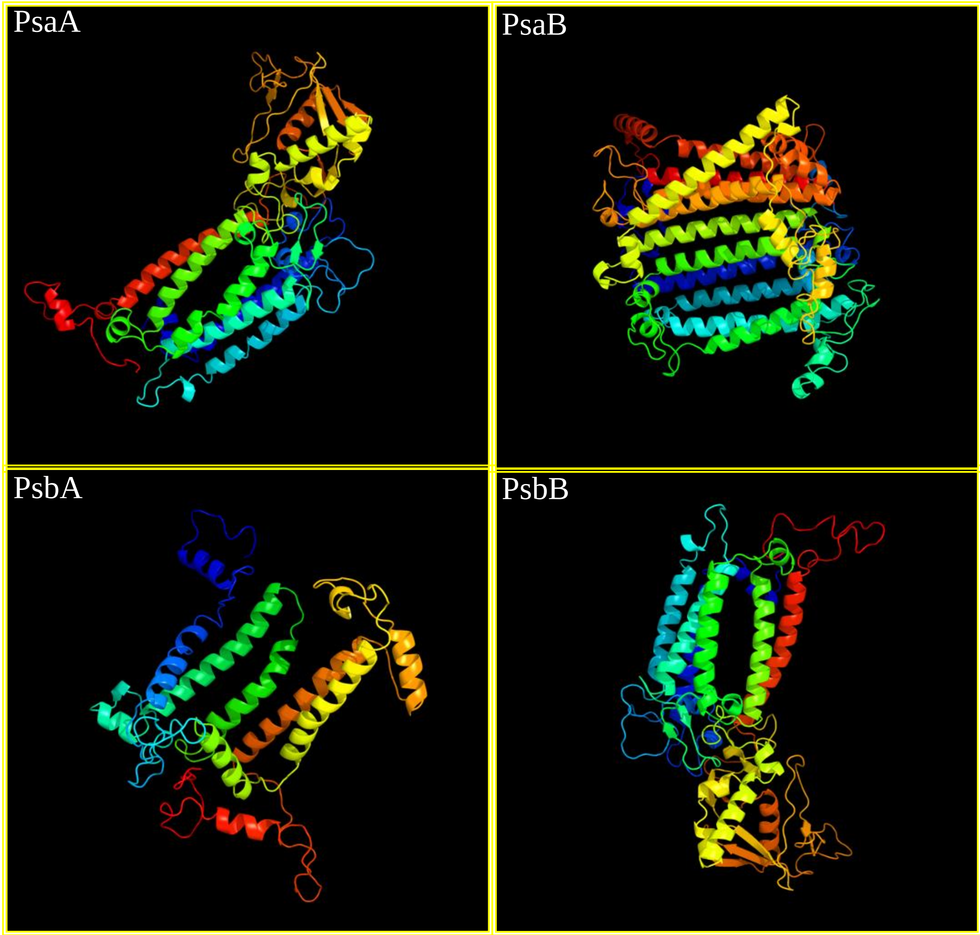

Fig. (8): Theoretical three-dimensional structure modeling of the PsaA (accession no. AML03238), PsaB (accession no. AML03239), PsbA (accession no. AML03240) and PsbB (accession no. AML03241) deduced amino acids sequences of *C. procera*. The three-dimensional structure models were constructed using Phyre^2 program (http://www.sbg.bio.ic.ac.uk/phyre2/).
Image were colored by rainbow N → C terminus.

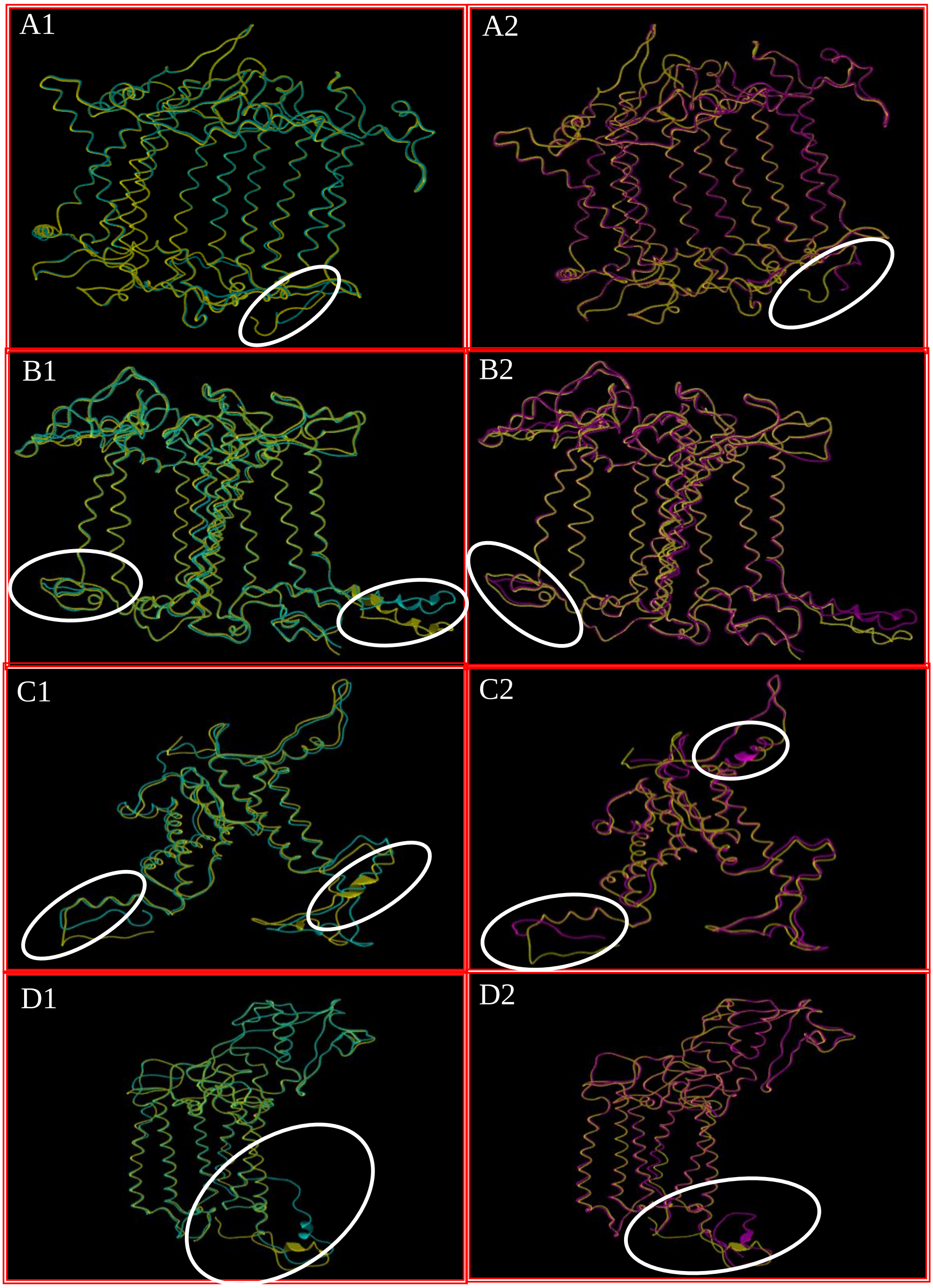

Fig. (9): Superimposition between PsaA, PsaB, PsbA and PsbB deduced amino acids sequences of *Arabidopsis thaliana* (blue, accession No. NP_051059.1, NP_051058.1, NP_051039.1, NP_051084.1, respectively) and *Glycine max* (purple, accession No. YP_538755.1, YP_538756.1, AAQ67339.1, YP_538791.1, respectively) and those of *C. procera* (yellow, accession No. AML03238, AML03239, AML03240, AML03241, respectively). A1 & A2: PsaA; B1 & B2: PsaB; C1 & C2: PsbA and D1 & D2: PsbB. TM-align program (http://zhanglab.ccmb.med.umich.edu/TM-align/).

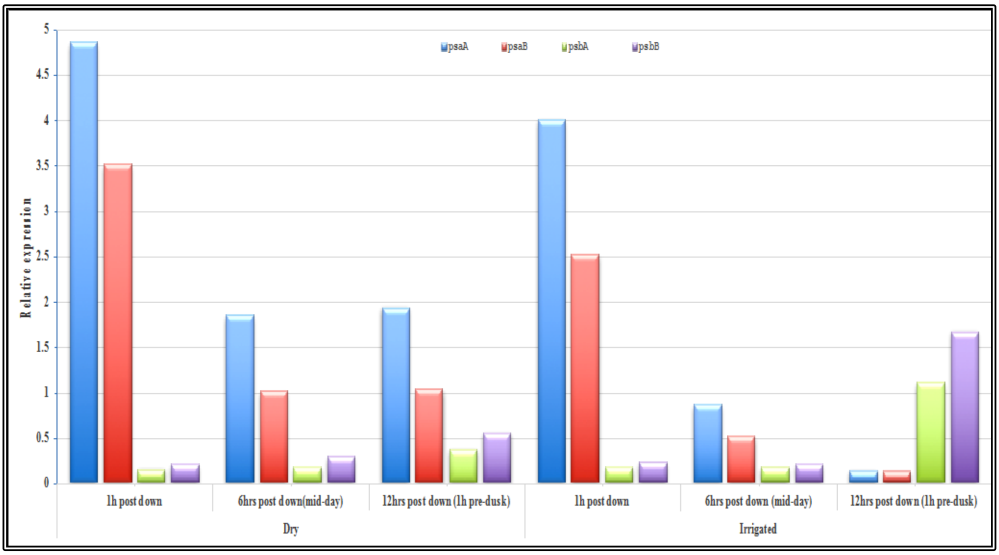

Fig. (10): Relative gene expression analysis of *psaA*, *psaB*, *psbA* and *psbB* transcripts of *C. procera* under sudden water availability. Dry: desert grown plants; irrigated: each plant receives 25 L water.