Arch. Biol. Sci., Belgrade, 64 (1), 65-70, 2012 DOI:10.2298/ABS1201065G

COMPUTATIONAL IDENTIFICATION OF miRNAs AND THEIR TARGETS FROM *CROCUS SATIVUS* L.

PRAVEEN GULERIA, DEEPMALA GOSWAMI and SUDESH KUMAR YADAV*

Plant Metabolic Engineering Laboratory, Biotechnology Division, CSIR-Institute of Himalayan Bioresource Technology, Council of Scientific and Industrial Research, Palampur-176061 (HP), India

Abstract - miRNAs are smaller known entities that perform several crucial regulatory roles in plants. Identification and characterization of miRNAs has been a challenging task which has become simplified with the development of computational approaches. In the present study, three novel miRNAs were predicted from *Crocus sativus* by computational approaches. A total 6767 ESTs were mined from the EST database available at NCBI. The software MicroPC was utilized which resulted with the prediction of three miRNAs, csa-miR1, csa-miR2 and csa-miR3. Targets were predicted for the respective miRNAs using the miRU2 program. The identified targets have a role in plant growth, senescence, disease resistance and various stress responses. In addition, some targets are involved in mRNA export, protein synthesis and post-translational modifications.

Key words: microRNA, *Crocus sativus*, structure, targets

INTRODUCTION

Micro RNAs (miRNAs) are small non-coding RNAs varying from 19-22 nucleotides (nts). They have been found to play major regulatory roles in plants, animals and fungi (Bartel and Bartel, 2003). miRNAs are endogenous in origin and are encoded by the genes present in the host's genome (Zeng et al., 2003). These miRNA-expressing genes can be traced onto the genome by various approaches such as genetic screening, direct cloning or computational predictions. Due to the non-availability of genomic sequence data and the tissue and time bound expressions of certain miRNAs, it is difficult to study them using the former two methods (Zhang et al., 2006). The computational approach involves the prediction of miRNAs from EST or GSS analysis. This approach has been determined as being cost effective and a rapid route towards the discovery and isolation of miRNAs. Originally, ESTs were used for the identification of genes involved in specific plant metabolic pathways. Recently ESTs have also been used for *in silico* identification of miRNAs from plants as well as animals (Zhang et al., 2005). Computational identification of miRNAs is based on certain guidelines. A predicted miRNA should possess the following features in order to be considered as a potent miRNA: (a) it should be 19-22 nt long (Bartel, 2004); (b) should have a well-predicted stem loop structure with low free energy (Zuker et al., 2003); (c) should not possess any loop or break in the predicted miRNA sequence; (d) usually mature miRNA for certain specific functions are conserved among plants (Xie et al., 2007; Sunkar and Jagadeeswaran, 2008).

Crocus sativus, commonly known as saffron, is a member of the family Iridaceae. Among the other 80 known species, this species of crocus is important because of the spice saffron which is derived from its flower. It is widely used in cooking as a season-

ing and coloring agent. Saffron is the world's most expensive spice by weight. Saffron has a bitter taste and a carotenoid dye, crocin, which allows it to impart a rich golden-yellow hue to dishes and textiles. Medicinally, saffron has been discovered to have anticarcinogenic, antimutagenic, immunomodulating and various antioxidant properties. It has also been used in cosmetics due to its property of reducing skin pigmentation (Sampathu et al., 1984). To date, no miRNA has been predicted from saffron crocus. In this study, an *in silico* method of EST analysis has been utilized for the prediction of miRNAs from *Crocus sativus* L. The predicted potential miRNAs were further analyzed for valid secondary structures using UNAFold software. Targets were predicted for these putative miRNAs that give us insight into their possible regulatory roles in *Crocus sativus*.

MATERIALS AND METHODS

Collection of ESTs

The ESTs of *Crocus sativus* were collected from dbEST available at NCBI (http://www.ncbi.nih.gov/dbEST/). A total 6767 ESTs were retrieved.

miRNA prediction

The application μPC (http://www.biotec.or.th/isl/micropc) was used for miRNA prediction. This software compares the query sequence against the known miRNAs from miRBase. *Crocus sativus* ESTs were submitted as the input query sequence with the condition <3000 bp. Default parameters were set. Results were obtained as possible complementary miRNAs to the input query sequence. The output shows conserved as well as novel miRNAs in the respective plant ESTs (Wuttichai and Duangdao, 2009).

Secondary structure prediction

The secondary structure of the miRNA within the EST was predicted with UNAFold using a "fold secondary structure" option. The substructure of stem-loop with a mature miRNA sequence was also predicted by specifying the start and stop position. MFE and MFEI values were obtained with the results as described earlier (Bhardwaj et al., 2010).

Prediction of targets for identified miRNAs

It has been reported that most of the known plant miRNAs regulate their mRNA targets by binding to the protein-coding region with perfect or near perfect complementarities (Wang et al., 2004). The potential targets for the predicted plant miRNAs were found using miRNA target finder program miRU which is available at http://bioinfo3.noble.org/miRNA/miRU (Zhang, 2005). *Arabidopsis thaliana* genome sequences were used as a base to predict the targets. The targets were identified on the basis of the complementarity of the identified miRNAs. Two criteria considered were no gap and <4 mismatches between miRNA and targets.

RESULTS AND DISCUSSION

Identification of potential miRNAs from Crocus sativus

Computational approaches have been employed for the miRNA prediction from several plant and animal systems (Qiu et al., 2007; Qiang et al., 2007; Takane et al., 2010). In the present work, 6767 ESTs was analyzed for miRNA prediction in the saffron crocus (*Crocus sativus* L.). The ESTs were retrieved from the NCBI EST database and were submitted to μPC. It has been demonstrated that a less than 4 nt mismatch is essential for the miRNA sequences to be homologous (Zhang, 2005). Depending on the criteria used for miRNA prediction from the ESTs of *Crocus*, none was found to be conserved. Only three novel miRNAs were predicted. The predicted miRNAs were named: csa-miR1, csa-miR2 and csa-miR3. These were 21-24 nt long. The majority of miRNAs from other plants also possessed the same size range (Xie et al., 2007; Sunkar and Jagadeeswaran, 2008). The A+U content of the predicted miRNAs varied from 40 to 86%. The newly identified miRNAs had the highest minimal fold energies (MFEs). Another useful criterion for distinguishing miRNAs from other types of coding

Table 1. List of miRNAs predicted from ESTs analysis of *Crocus sativus*.

miRNA	EST ID	Sequence	End	A+U (%)	MEFI
csa-miR1	157008739 157007541 157008628 157007216 157006740 157006506 157006497 157008521 157008368 157007794 157010423 157011146	AUAAGUGGGUUUGUGUUAAUAUGA (24 nt)	3’	70	0.83
csa-miR2	157008100 157007625 157006384 157011159 157010624 157010459	UCCCUCUCAUCAUCCUCGUCG (21 nt)	3’	42.84	0.61
csa-miR3	157007825 157007325 157006654	AAUUAUAAUUAAGUUUUUAGC (21nt)	3’	85.68	0.73

Table 2. List of targets identified for the novel miRNAs predicted for *Crocus sativus*.

miRNA	Target Accession ID	Target sequence **(5’ - 3’)**	Protein
csa-miR1	At4g33950.1	UUAUAUUAACACCAACCCAUUUAG	abscisic acid-activated protein kinase
csa-miR2	At1g20920.1	AGAUGAUGAUGAAGAGAGGGA	DEAD box RNA helicase
	At3g47860.1	UGAUGAUGAUGAUGAGAGGCA	Apolipoprotein D-related Leucine-rich repeat family protein
	At1g49490.1	CGACGAUGAUGAUGAUGGGGA	hAT dimerisation domain-containing protein
	At5g33406.1	UGAUGAGGAUGAAGAGAUGGA	Peptide chain release factor
	At2g47020.1	AGAUGAAGCUGAUGAGAGGGA	
csa-miR3	At2g25690.1	CCUAAAGACUUUAUUAUAAUA	Senescence-associated protein

and non-coding RNAs is the minimal free energy index (MFEI). The precursors of the miRNAs had higher MFEIs than other types of RNAs. The MFEIs of the predicted miRNAs were in the range of 0.63-1.32 (Table 1). The length of pre-miRNAs varied from 101-151 bp. These parameters were in agreement with previously reported results for *in silico* predicted miRNAs from various plants (Qiu et al., 2007; Qiang et al., 2007; Takane et al., 2010).

miRNAs are known to differ from other RNAs due to the distinguished property of their precursor sequences to form a secondary hairpin structure (Bhardwaj et al., 2010; Zhang, 2005). Therefore, the secondary structures of all the identified miRNAs were predicted using UNAFold (Fig. 1). The predicted secondary structure of the identified miRNAs provides further confirmation for considering them as miRNAs. The identified miRNAs were found to

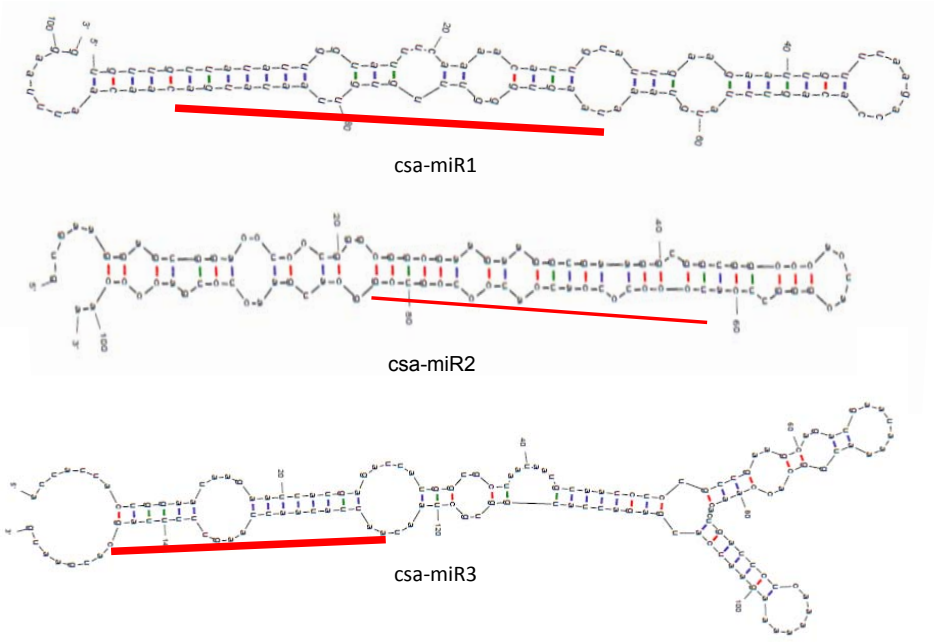

Fig. 1. Predicted secondary structure of identified precursor miRNAs in *Crocus sativus* L. These structures were produced using the UNAFold program. The mature miRNAs sequences are marked with a line. The actual size of the precursors may be slightly shorter or longer than those presented here.

be located at 3' ends of their respective precursor sequences.

Prediction of targets for newly identified miRNAs and their putative role

Insights into miRNA targets can help in understanding their role and functional importance. The predicted targets for the identified miRNAs are shown in Table 2. miRU software was employed to predict miRNA targets. Most of the predicted targets were found to play a crucial role in the growth, development and physiological responses of the plants. Among the predicted three miRNAs, csa-miR2 was found to target five regulatory genes. Regulatory genes as targets are an important feature of miRNAs (Zhang et al., 2006).

csa-miR1 was found to target the gene encoding the abscisic acid-activated protein kinase. It has been reported that ABA-mediated protein kinases play a crucial role in plant response to water stress. ABA-mediated protein kinase mutants of *Arabidopsis* have shown a loss of stomatal movements and unresponsiveness to humidity and dehydration stress (Yoshida et al., 2002). In addition, several studies have elucidated the role of miRNAs in plant development and stress response (Phillips et al., 2007). csa-miR2 was found to target more than one regulatory gene. Among the predicted targets, two Apolipoprotein D-related and DEAD Box RNA Helicase were identified as stress response regulatory proteins. *Arabidopsis* apolipoprotein D ortholog has been demonstrated to possess a role in modulating oxidative stress response (Charron et al., 2008). Similarly, DEAD Box RNA helicase is involved in mRNA export, plant development and temperature stress response (Gong et al., 2005). A leucine-rich repeat (LRR) protein is the other target predicted for csa-miR2. These repeat proteins have been found to regulate plant pathogen resistance genes via signal transduction. LRRs play a role in pathogen ligand recognition and hence undergo disease management (Shanmugam, 2005). The rest of the targets were observed to participate in pro-

tein synthesis and modifications. One of the target hAT dimerization domain-containing proteins has a role in protein dimerization while the target peptide chain release factors are required for the termination of protein synthesis.

Senescence is one of the crucial aspects of plant development, and miRNAs are also known to target senescence responsive genes (Schommer et al., 2008). csa-miR3 was found to target such senescence-associated protein encoding genes. Senescence-associated proteins are primarily expressed in the senescing tissues of plants. One of the senescence-associated genes from *Arabidopsis* has been reported to regulate cellular viability during salinity and osmotic stress conditions (Seo et al., 2010). The predicted potent miRNAs of *Crocus sativus* have been found to target various growth- and development-related regulatory genes.

CONCLUSION

This study presents the computational miRNA prediction of *Crocus sativus* via EST analysis. The targets of predicted miRNAs assume roles in plant growth, defence, stress response and senescence. In addition, they participate in protein synthesis and modifications. The predicted miRNAs do not show significant homology with any of the previously identified and characterized miRNAs. Therefore, the predicted miRNAs can be considered to be novel molecules and can be grouped as members of a new family.

Acknowledgements – The authors are grateful to the Director IHBT for his continuous support and encouragement to pursue research in this direction. The authors would like to thank the Council of Scientific and Industrial Research (CSIR), Govt. of India for financial support. PG would like to thank CSIR for providing fellowship in the form of Junior Research Fellowship (JRF).

REFERENCES

Bartel, B., and *D.P. Bartel* (2003). MicroRNAs: At the root of plant development? *Plant Physiol.* **132**, 709-717.

Bartel, D. P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297.

Bhardwaj, J., Mohammad, H., and *S. K. Yadav* (2010). Computational identification of MicroRNAs and their targets from the expressed sequence tags of horsegram (*Macrotyloma uniflorum* (Lam.) Verdc.). *J. Struct. Funct. Genomics* **11**, 233-240.

Charron, J. F., Ouellet, F., Houde, M., and *F. Sarhan* (2008). The plant Apolipoprotein D ortholog protects *Arabidopsis* against oxidative stress. *BMC Plant Biol.* **8**, 86-98.

Gong, Z., Dong, C. H., Lee, H., Zhu, J., Xiong, L., Gong, D., Stevenson, B., and *J. K. Zhu* (2005). A DEAD Box RNA Helicase Is Essential for mRNA export and important for development and stress responses in *Arabidopsis*. *Plant Cell* **17**, 256-267.

Phillips, J. R., Dalmay, T., and *D. Bartels* (2007). The role of small RNAs in abiotic stress. *FEBS Lett.* **581**, 3592-3597.

Qiang, G., AnLing, X., Qing, Y., and *Y. ZhiMin* (2007). Bioinformatic identification of microRNAs and their target genes from *Solanum tuberosum* expressed sequence tags. *Chinese Science Bulletin* **52**, 2380-2389.

Qiu, C. X., Xie, F. L., Zhu, Y. Y., Guo, K., Huang, S. Q., Nie, L., and *Z. M. Yang* (2007). Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. *Gene* **395**, 49-61.

Sampathu, S. R., Shivashankar, S., Lewis, Y. S., and *A. B. Wood* (1984). Saffron (*Crocus Sativus* Linn.)- Cultivation, processing, chemistry and standardization. *CRC* **20**, 123-157.

Schommer, C., Palatnik, J. F., Aggarwal, P., Chetelat, A., Cubas, P., Farmer, E. E., Nath, H., and *D. Weigel* (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. *PLoS Biol.* **6**, e230.

Seo, P. J., Park, J. M., Kang, S. K., Kim, S. G., and *C. M. Park* (2011). An *Arabidopsis* senescence-associated protein SAG29 regulates cell viability under high salinity. *Planta* **233**, 189-200.

Shanmugam, V. (2005). Role of extracytoplasmic leucine rich repeat proteins in plant defence mechanisms. *Microbiological Res.* **160**, 83-94.

Sunkar, R., and *G. Jagadeeswaran* (2008). *In silico* identification of conserved microRNAs in large number of diverse plant species. *BMC Plant Biol.* **8**, 37-49.

Takane, K., Fujishima, K., Watanabe, Y., Sato, A., Saito, N., Tomita, M., and *A. Kanai* (2010). Computational prediction and experimental validation of evolutionarily conserved microRNA target genes in bilaterian animals. *BMC Genomics* **11**, 101-113.

Wang, X. J., Reyes, J. L., Chua, N. H., and *T. Gaasterland* (2004). Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol.* **5**, R65.

Wuttichai, M. and *W. Duangdao* (2009). MicroPC (μPC): A comprehensive resource for predicting and comparing plant MicroRNAs. *BMC Genomics* **10**, 366-373.

Xie, F. L., Huang S. Q., Guo, K., Xiang, A. L., Zhu, Y. Y., Nie, L., and *Z. M. Yang* (2007). Computational identification of novel microRNAs and targets in *Brassica napus*. *FEBS Lett.* **581**, 1464-1474.

Yoshida, R., Hobo, T. Ichimura, K., Mizoguchi, T., Takahashi, F., Aronso, J., Ecker, J. R., and *K. Shinozaki* (2002). ABA-activated SnRK2 protein kinase is required for dehydration stress signaling in Arabidopsis. *Plant Cell Physiol*. **43**, 1473-1483.

Zeng, Y., Yi, R. and *B. R. Cullen* (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl. Acad. Sci. USA* **100**, 9779-9784.

Zhang, B. H., Pan, X. P., Wang, Q. L., Cobb, G. P., and *T. A. Anderson* (2006). Computational identification of microRNAs and their targets. *Comput. Biol. Chem.* **30**, 395-407.

Zhang, B. H., Pan, X. P., Wang, Q. L., Cobb, G. P., and *T. A. Anderson* (2005). Identification and characterization of new plant microRNAs using EST analysis. *Cell Res.* **15**, 336-360.

Zhang, Y. (2005). miRU: An automated plant miRNA target prediction server. *Nucleic Acids Res.* **33**, W701-W704.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406-3415.