J. Agr. Sci. Tech. (2015) Vol. 17: 1637-1649

Molecular Cloning and Characterization of a *Cyclotide* Gene Family in *Viola modesta* Fenzl

N. Kodari[1], B. Bahramnejad[1*], J. Rostamzadeh[2], H. Maroofi[3], and S. Torkaman[1]

ABSTRACT

Cyclotides are small disulfide-rich proteins that have the unusual feature of a cyclic backbone. Cyclotides have a range of interesting biological activities and are found in a variety of tropical plants from the *Rubiaceae, Violaceae, Cucurbitaceae* and *Fabaceae* families. We have cloned and characterized cyclotides in *Viola modesta*, a *Viola* species native to western Asia, which was collected from the Kurdistan Province of Iran. Fifteen cyclotide sequences were obtained using homology based PCR strategy. Sequence analysis showed that 14 of them had continued open reading frames and showed high level of similarity to cyclotide *genes* from other species of the *Violaceae*. After analyzing the full endoplasmic reticulum signals of *V. modesta* cyclotides, two conserved sequences, AAFALPA and ATAFALP, were detected. Analysis of isolated cyclotide sequences showed that they all belonged to bracelet family and were separated into two subclasses. Phylogenetic analysis of *cyclotide* genes from *V. modesta* and other *Viola* species revealed that most *V. modesta* genes showed close relationship with their homologs from the *Violaceae*, while the *V. modesta* genes formed two separate clades. Transcription analysis by semi-quantitative RT-PCR revealed that *Vmcyc1* and *Vmcyc7* were differentially expressed in all tested tissues including roots, stems, leaves, flowers, seeds, peduncles, and capsules with the highest transcript level in the capsules.

Keywords: Bracelet cyclotides, Phylogeny, RT-PCR, Sequences, 3'RACE.

INTRODUCTION

Cyclotides are a family of circular plant proteins containing approximately 30 amino acids (Craik *et al*., 1999). Their *N*-terminal and *C*-terminal are connected via an amide bond that results in a circular peptide backbone. Cyclotides have six cysteine residues that each two cysteines linked by a disulphide bond. Combining three disulphide bonds and cyclic structure define the Cyclic Cysteine Knot (CCK) motif (Herrmann *et al*., 2006). The CCK motif is responsible for the cyclotides extraordinary stability against chemical and thermal degradation (Herrmann *et al*., 2008). Cyclotides are responsible for a range of biological activities including anti-HIV, uterotonic, antimicrobial, hemolytic, cytotoxic, neurotensin antagonistic, antifouling, and pesticide (Craik, 2012).

The cyclotide proteins fall into two major families including Mobius and bracelet. The Mobius cyclotides contain a *cis*-pro peptide bond in loop 5, which creates a twist in the peptide backbone (Herrmann *et al*., 2006). Most Mobius cyclotides are slightly negatively charged or have an overall net-charge of zero, while bracelet cyclotides are usually positively charged. Bracelet cyclotides are more

[1] Department of Agricultural Biotechnology, Faculty of Agriculture, University of Kurdistan, Sanandaj, Islamic Republic of Iran.
* Corresponding author; e-mail: b.bahramnejad@uok.ac.ir
[2] Department of Animal Science, Faculty of Agriculture, University of Kurdistan, Sanandaj, Islamic Republic of Iran.
[3] Research Center of Agriculture and Natural Resources, Kurdistan Province, Sanandaj, Islamic Republic of Iran.

abundant, since, until now, the majority (~70%) of more than 200 published cyclotide sequences accessible on CyBase (Wang *et al*., 2008) belong to this subfamily.

Analysis of cyclotide precursor sequences show that genes encoding them consist of an Endoplasmic Reticulum (ER) signal domain, a pro-region and one to three mature cyclotide domains, each preceded by a *N*-Terminal Repeat (NTR) sequence (Zhang *et al*., 2009). The NTR and cyclotide regions can be repeated up to three times in different precursors, encoding different or identical cyclotides (Burman *et al*., 2010). It has been suggested that an asparaginyl-endoproteinase is involved in cleavage of the *C*-terminal tail and simultaneous cyclization of the cyclotide. Details of the processing of the precursors, including the order of the events, are not fully understood, but they mostly mediate through the oxidative folding, excision of the mature cyclotide sequence, and head-to-tail cyclization.

More than 198 cyclotides have been discovered from 36 species in the *Violaceae, Rubiaceae, Cucurbitaceae, and Fabaceae* and *cyclotide-like* genes in Poaceae plant families (Gerlach and Mondal, 2012). Most of sequenced cyclotides have been isolated from species of the *Violaceae*. The *Violaceae* family contains around 23 genera and 800 species of cosmopolitan shrubs, herbs, and rare trees (Gerlach and Mondal, 2012). Within the *Violaceae*, cyclotides seem to be widely distributed, but the cyclotide complements of the vast majority of *Violaceae* species have not yet been fully explored. *Cyclotide* genes have been isolated from *V. arvensis* (Mulvenna *et al*., 2005), *V. odorata* (Dutton *et al*., 2004), *V. biflora* (Herrmann *et al*., 2008), *Hybantus floribundus* (Simonsen *et al*., 2005) and *V. baoshanensis* (Zhang *et al*., 2009).

Viola modesta Fenzl, also known as *V. ebracteolata* Fenzl or *V. modestula* Klokov, is an annual viola species native to Iran, Iraq, Israel, Jordan, Lebanon, Syria, Turkey and Turkmenistan. *V. modesta* is a diploid plant with chromosome number (2n=2x=4) in an annual species with the lowest chromosome number so far known for a vascular plant (Erben, 1996). *V. modesta* is mainly distributed in the west of Iran, Kurdistan Province. It is naturally rare and has a patchy distribution.

In the present study, we aimed to describe the *cyclotide* genes isolated via homology based PCR strategy from *V. modesta*, and their expression patterns in various tissues.

MATERIALS AND METHODS

Plant Materials

Viola modesta Fenzl was collected from Greize Research Station, Sanandaj, Iran, in May 2010. Plant organs including roots, stems, leaves, flowers, seeds, peduncles, and capsules were harvested and immediately immersed in liquid nitrogen for RNA extraction.

Extraction of RNA and cDNA Synthesis

RNA was extracted from different tissues via Mazzara Protocol (Mazzara and James, 2000). The concentration and integrity of isolated RNA were determined by spectrophotometer and 1.5% agarose gel electrophoresis and stored at −70°C. The first-strand cDNA was synthesized with 4 µg total RNA in 20 µL reaction volume using first strand cDNA synthesis Kit according to the manufacturer's instruction (Fermentas).Total RNA (1 µg) was treated by DNase1, RNase-free kit (Fermentas). After DNA digestion, the RNA samples were first heated (65°C, 5 minutes) to inactivate the DNase I and avoid RNA secondary structures, and then immediately cooled in ice water.

Isolation of 3′ end by RACE

An alignment using CLUSTALW was made of *cyclotide* genes from *Viola biflora*; EU046618.1, EU046619.1, EU046621.1, EU046622.1, EU046623.1, that were obtained

from NCBI. A forward primer was designed from conserved regions AAFALPA, SF1, 5' *GATTCCAAGATTGTGTTTGTAGCC*. Another primer, *ATGAAGATGTTTGTTGCCCTT*, was designed based on *Mobius* genes, Vbc.7c, Vbc6c, Vbc6d, Vbc6, Vbc7, Vbc7b and Vbc5 From *V. baoshanensis* to amplify a second class of cyclotides in *V. modesta*. First strand cDNA for 3′ RACE was synthesized through reverse transcription with 3′ CDS (5′ *GACCACGCGTATCGATGTCGACTTTTTTT TTTTTTTTV* 3′) as anchor primer using first strand cDNA synthesis Kit. The first round PCR reaction was carried out in a total volume of 25 µL including 8 µL H_2O, 12 Master Mix, 2 µL First strand cDNA template, 1.5 µL SF1, 1.5 µL PCR anchor primer (GACCACGCGTATCGATGTCGAC). PCR temperature program was 1 cycle of 5 minutes at 94°C, 35 cycles of 30 seconds at 94°C, 30 seconds at 53°C, 40 seconds at 72°C, 1 cycle of 20 minutes at 72°C. The resulting PCR product was separated on 1% agarose gels.

Molecular Cloning and DNA Sequencing

PCR products were separated on 1.0% agarose gels and the expected fragments were purified from the gels using Nucleic Acid Extraction kit (Vivantis). The concentration of purified DNA was determined by spectrophotometer, then, DNA fragments were ligated into the TA vector using TA cloning kit (Fermentas) and transformed into competent cells of *Escherichia coli* DH5a strain. Positive clones were identified by colony PCR, and independent sequences per clone were obtained from a commercial sequencing service (Bioneer Inc. Bioneer Corporation).

Bioinformatics Analysis

BLAST program in National Center for Biotechnology Information Server (http://www.ncbi.nlm.nih.gov.) was used to verify the *cyclotide* gene homology. Multiple protein sequence alignment was performed using Clustal W program of Jalview 2.3 version. For phylogenetic and sequence alignment analysis, other *cyclotide* gene sequences were obtained from the GenBank database (Table 1). The phylogenetic tree of cyclotides was constructed using MEGA4.0.2 software based on the Neighbor-Joining (NJ) Method. Theoretical isoelectric point and mass values for the protein was predicted using ExPASy ProtParam tool (http://us.expasy.org.tools.protparam.html). The 3D structures of two cyclotide *Vmcyc1* and *Vbc1* (*Viola biflora* cyclotide 1) were obtained using internet http://swissmodel.expasy.org. The ER signals were predicted using signalP3.0 (Bendtsen *et al*., 2004).

Gene Expression Analysis

For gene expression analysis, roots, stems, leaves, flowers, seeds, peduncles, and capsules of *V. modesta* were harvested and immediately immersed in liquid nitrogen for RNA extraction. Total RNA was extracted and treated with DNase 1 (DNaseI, RNase-free Fermentas) to exclude the DNA contamination. The first-strand cDNA used as template was synthesized with 4 µg total RNA in 20 µL reaction volume using first strand cDNA synthesis Kit (Fermentas). *Cyclotide* genes obtained from *V. modesta* were divided into cluster. For expression analysis, *Vmcyc1* and *Vmcyc2* were selected and specific primer pairs were designed using primer blast. *Vmcyc 1* gene fragments was amplified using specific forward primer VmcycF1 (5′GCAGTACTCCTTGTGGAGAGA3′), and antisense primer Vmcyc R1 (5'AGACAAAGACAACTTCTTCCACA3'). *Vmcyc*7 fragments was also amplified using specific sense forward primer VmcycF7 (5'ATCTGCACCACACTCAAACATC3')

Table 1. Cyclotide amino acid sequences used in this study.

Total Nucleotide (bp)	Molecular weight (Kd)	Isoelecteric point	Total Amino acid	Cyclotide group	Accession Number[a]	Source	Cyclotide name
568	11.025	8/29	104	Bracelet	Current study	*V. modesta*	Vmcyc16
592	10.56	4/91	104	Bracelet	Current study	*V. modesta*	Vmcyc1
573	10.81	6/08	104	Bracelet	Current study	*V. modesta*	Vmcyc8
570	10.81	6/08	104	Bracelet	Current study	*V. modesta*	Vmcyc17
574	10.82	5/08	104	Bracelet	Current study	*V. modesta*	Vmcyc4
581	10.84	6/08	104	Bracelet	Current study	*V. modesta*	Vmcyc3
585	10.86	6/08	104	Bracelet	Current study	*V. modesta*	Vmcyc9
573	10.78	6/08	104	Bracelet	Current study	*V. modesta*	Vmcyc12
642	12.48	6/05	118	Bracelet	AY630564/1	*V. odorata*	Voc. C1
615	12.41	6/05	118	Bracelet	AY630563/1	*V. odorata*	Voc.C2
527	12.15	5/30	113	Bracelet	EU910536/1	*V. baoshanensis*	Vbc.2b
676	12.15	5/30	113	Bracelet	DQ85186/1	*V. baoshanensis*	Vbc.2
528	12.07	5/30	113	Bracelet	EU910537/1	*V. baoshanensis*	Vbc.2c
528	12.14	6/53	113	Bracelet	EU910535/1	*V. baoshanensis*	Vbc.2a
528	12.01	4/96	113	Bracelet	EU910538/1	*V. baoshanensis*	Vbc.2d
409	12.07	5/71	113	Bracelet	EU910534/1	*V. baoshanensis*	Vbc.1c
409	12.16	5/71	113	Bracelet	EU910553/1	*V. baoshanensis*	Vbc.1b
643	12.16	5/71	113	Bracelet	DQ851860/1	*V. baoshanensis*	Vbc.1
409	12.13	5/71	113	Bracelet	EU910532/1	*V. baoshanensis*	Vbc.1a
567	12.50	5/76	115	Bracelet	EU910542/1	*V. baoshanensis*	Vbc.3c
686	12.50	5/76	115	Bracelet	DQ851862/1	*V. baoshanensis*	Vbc.3
550	12.10	7/55	111	Bracelet	EU910540/1	*V. baoshanensis*	Vbc.3a
566	12.62	8/22	116	Bracelet	EU910541/1	*V. baoshanensis*	Vbc.3b
521	12.44	6/21	115	Bracelet	EU910543/1	*V. baoshanensis*	Vbc.3d
545	12.44	6/21	115	Bracelet	AY630565/1	*V. odorata*	Voc.C3
499	12.23	7/74	111	Bracelet	Current study	*. modesta*	Vmcyc40
459	10.50	6/43	98	Bracelet	Current study	*. modesta*	Vmcyc5
502	10.68	5/19	100	Bracelet	Current study	*. modesta*	Vmcyc70
465	10.53	5/19	100	Bracelet	Current study	*. modesta*	Vmcyc11
479	10.52	6/78	98	Bracelet	Current study	*V. modesta*	Vmcyc7
556	12.83	8/22	120	Bracelet	EU910554/1	*V. baoshanensis*	Vbc.4a
726	12.83	8/22	120	Bracelet	DQ851863/1	*V. baoshanensis*	Vbc.4
546	12.87	8/22	120	Bracelet	EU910546/1	*V. baoshanensis*	Vbc.4c
546	12.82	6/54	120	Bracelet	EU910545/1	*V. baoshanensis*	Vbc.4b
485	11.30	6/55	104	Mobius	EU910547/1	*V. baoshanensis*	Vbc.5
840	21.33	5/88	207	Mobius	AY630566/1	*V. odorata*	Voc.k1
624	21.33	5/88	207	Mobius	FJ211181/1	*V. odorata*	Voc.O8
606	15.76	6/21	153	Mobius	EU910549/1	*V. baoshanensis*	Vbc.6c
769	21.38	7/31	207	Mobius	EU910552/1	*V. baoshanensis*	Vbc.7c
768	21.55	6/54	207	Mobius	EU910551/1	*V. baoshanensis*	Vbc.7b
768	21.26	6/28	207	Mobius	EF583937/1	*V. baoshanensis*	Vbc.7
606	15.79	6/21	153	Mobius	EU910550/1	*V. baoshanensis*	Vbc.6d
606	15.78	6/21	153	Mobius	EF583936/1	*V. baoshanensis*	Vbc.6
549	11.07	6/77	105	Bracelet	EU046618/1	*V. biflora*	Vbi.1
535	10.95	5/72	103	Bracelet	EU046619/1	*V. biflora*	Vbi.2
596	11.11	6/78	105	Bracelet	AY630565/1	*V. biflora*	Vbi.3
553	10.97	5/72	103	Bracelet	EU046621/1	*V. biflora*	Vbi.4
529	11	5/72	103	Bracelet	EU046622/1	*V. biflora*	Vbi.5
542	10.92	5/72	103	Bracelet	EU046623/1	*V. biflora*	Vbi.6

and antisense primer VmcycF7 (5'GATAAGAAAGAGATGCGAATAGA3') . *Tubulin* gene fragment from *Viola cornuta* (GenBank accession number AY294027.1) was amplified as housekeeping gene for normalization using the sense primer TubF, (5'GAGGTTTGATGGAGCTCTTAATG3') and antisense primer TubR (5'GGTGGAATTGGAGATCATGCA3'). The PCR reaction system including 11.5 µL

H_2O, 12.5 µL 2×Master mix (Fermentas, USA), 1 µL sense primer, 1 µL antisense primer, 1 µL template cDNA. The PCR condition was as follows: 5 min pre-amplification at 94°C, 26 cycles of 30 seconds at 94°C, 40 seconds at 52°C, 1 minute and 30 seconds at 72°C, a final extension of 10 minutes at 72°C. The resulting PCR products were separated on 1% agarose gels. Independent experiment was repeated at least three times with the similar results.

RESULTS

Isolation and Sequence Analysis of *Vmcyc* Genes

Based on the combination of five *Viola biflora* Expressed Sequence Tags (ESTs), EU046618.1, EU046619.1, EU046621.1, EU046622.1, EU046623.1, and designing forward primers a band of the predicted size (~500 bp) was observed after PCR amplification with the cDNA synthesized from total RNA extracted from whole plant tissues as a template.

These expected fragments were excised from agarose gel and cloned into the plasmid vector (TA cloning, Fermentas). Positive clones were picked and used for screening. Twenty unique clones were chosen for DNA sequencing, of which 14 clones contained the primer sites. BLAST analysis against GenBank database revealed that the 14 genes were highly homologous to *cyclotide* genes (Table 1). Conceptual translations of the above 14 sequences revealed the presence of premature stop codons in six clones.

The sequences named as *Vmcyc* were regarded as *cyclotide* genes by the presence of continued Open Reading Frame (ORF) and by the characteristics of the motifs such as six cysteines and AAFALPA motif (Figure 1). For these 14 sequences, nucleotide identity among each clone pair was determined. The identity ranged from 54-99%, with the highest identity between *Vmcyc7*, *Vmcyc5,* and *Vmcyc10*, *Vmcyc70*. BLAST analysis of the *V. modesta cyclotide* genes against GenBank database revealed that they were homologous to *cyclotides* gene in *Violaceae* with *E*-values $5e^{-99}$< e×10.

Two types of conserved sequences were observed from the ER signals of *cyclotide* genes of *V. modesta*, Including ATFALPS (F) and AAFALPA (Figure 1). After sequence alignment and cluster analysis, 14 unique cyclotide sequences were identified and denoted *Vmcyc1*, 3, 4, 5, 7, 8, 9, 10, 11, 12, 16, 17, 40, and 70 (Table 1). *Vmcyc5*, 7, 40, 10, 11, and 70 starts with ATFALPAS (F) and mature cyclotide in these cyclotides start with GIP (Vmcyc5, 7, 40) and VNG (Vmcyc10, 11, 70), while *Vmcyc*17, 12, 9, 8, 16, 1, 4, 3 start with AAFALPA and mature cyclotide in these cyclotides start with GGT and GGS (just in Vmcyc1).

Most of *V. modesta* cyclotides were similar to *V. odorata* cyclotides (Voc1 and Voc3), except for *Vmcyc11* and *Vmcyc3* that were more similar to *V. biflora* cyclotide (vbc.3). The highest identity percentage was observed between *V. modesta* cyclotide *Vmcyc7* and *Vmcyc70* with *Vmcyc5* and *Vmcyc10* (99%), respectively. The smallest identity was observed between *Vmcyc40* with *Vmcyc4* and *Vmcyc1* (54%).

When nucleotide sequences were translated to amino acid sequences, different parts of cyclotide protein were defined (6 cycteines and AAFALPA sequence). Alignment of predicted cyclotide precursors from *V. modesta* showed that the cyclotide precursors had the overall arrangement in common with previously known cyclotide proteins from other *Violaceae* family (Figure 1). The cyclotide precursors consisted of an Endoplasmic Reticulum (ER) signal peptide which was 10 bp, similar to previously known cyclotides. *N*-Terminal ProPeptide (NTPP) was made of *N*-Terminal ProDomain (NTPD) and *N*-Terminal Repeat (NTR) and had 19-23 bp length. The mature cyclotide domain was 28-30 bp and a tail region in its C terminus 3-4 bp. The cystine knot motif present in the cyclotides coupled to the head to tail cyclic backbone is a

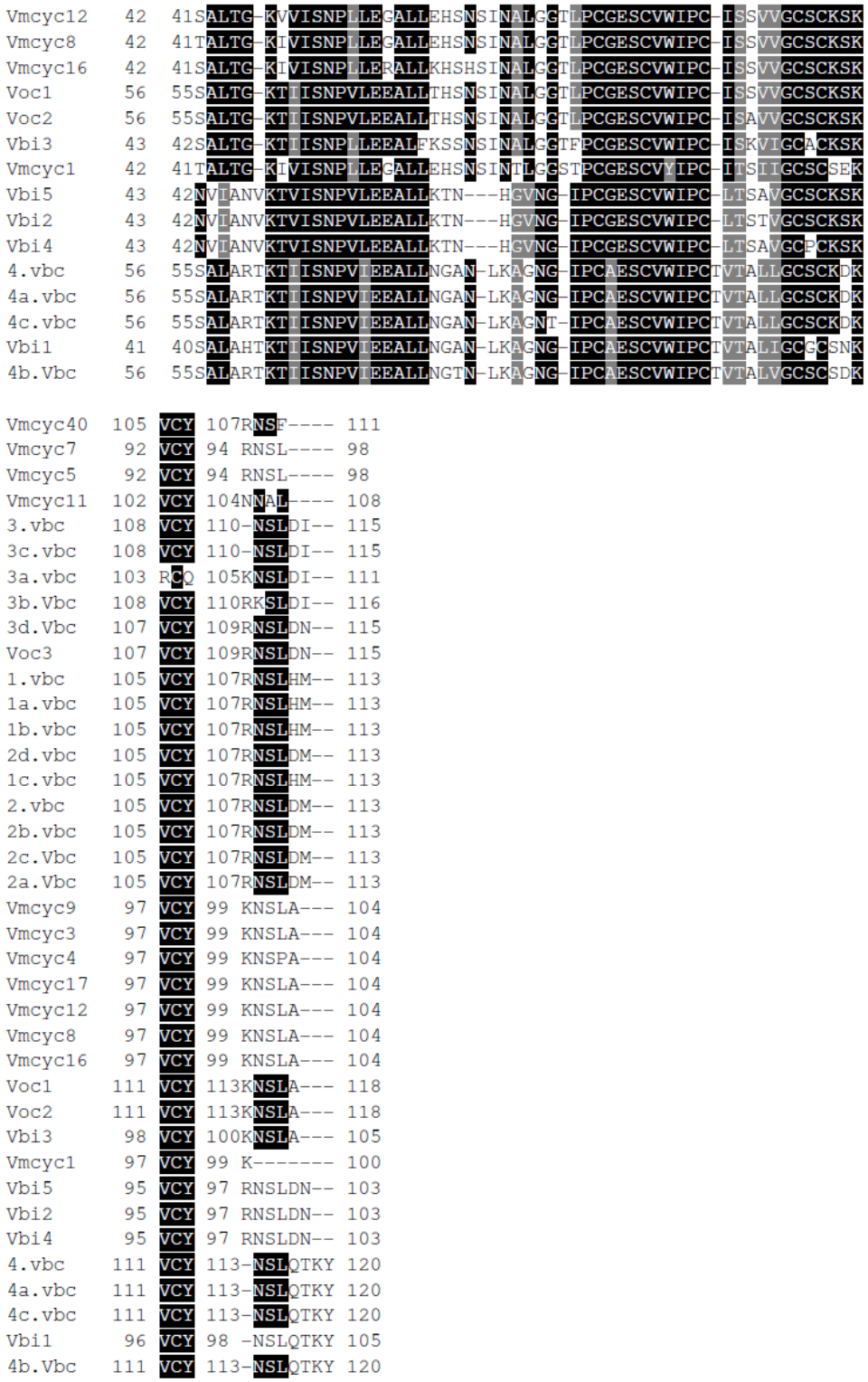

Figure 1. *Violaceae* cyclotides amino acid alignment, alignment of predicted cyclotide precursors from *V. modesta* and other viola species. These precursors have the overall arrangement in common with previously known CPs. ER signals were predicted using signalP3.0 (Bendtsen *et al*., 2004). The Genbank accession numbers of *Viola* species are in Table1.

unique motif called Cyclic Cystine Knot (CCK), which gives thermal, enzymatic, and chemical stability (Craik *et al*., 2001). The conserved residues across all bracelets are six Cys residues making up cystine knot and some residues. An efficient way to describe and compare the features of cyclotides is by referring to the intercysteine loops, illustrated in Figure 2 as an amino acid incidence plot for the 15 new *V. modesta* sequences in sequence logo format. Conserved residues among sequences include Pro3, Cys4, Glu6, Cys8, Val9, Pro1, Cys13, Gly19, Cys20, Ser21, Cys22, Lys25, Val26, Cys27, Tyr28, and Lys29. According to the sequence logo plot, the greatest variations in loop size and/or composition are in loops 3 and 6.

A ribbon representation of the structure of *Vmcyc11* (*Viola modesta* cyclotide 11) and *Vbc1* (*Viola baoshanensis* cyclotide 1) are shown in Figure 3. Both proteins had similar structures and were in good agreement with the previous Viola cyclotides in terms of number and position of beta sheets.

Phylogenetic Profile of *V. modesta Cyclotide* Genes

Phylogenetic profile of *V. modesta cyclotide* genes was carried out using MEGA software. The phylogenetic tree showed that *V. modesta cyclotide* genes were separated into two distinct clades, consistent with the similarity based amino-acid sequence analysis as described previously (data not shown). Two

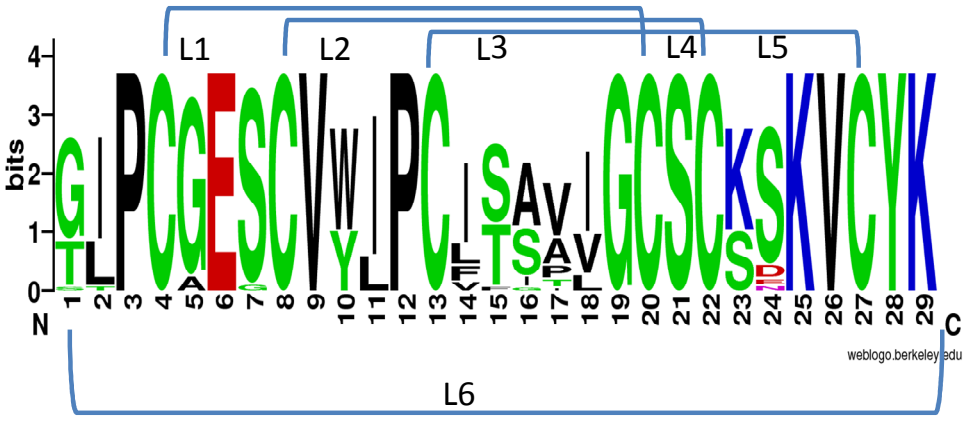

Figure 2. The pattern of conserved and variable parts as visualized by a sequence logo (*V. modesta* cyclotides and other database cyclotides). The cyclotide loops are shown (L1 to L6). The overall height of each stack of letters indicates the sequence conservation. While the height of each letter in a stack reflects the relative frequency of the corresponding amino acid.

Vmcyc11 Vbc1

Figure 3. *Vmcyc11* (*Viola modesta* cyclotide 11) and *Vbc1* (*Viola baoshanensis* cyclotide 1) structures. (http://swissmodel.expasy.org)

types of conserved sequences were observed from the ER signals of CPs, including ALVLIATFA and AAFALPA-LA. An evolutionary relationship among *V. modesta cyclotide* genes and other known *cyclotide* genes from *Violaceae* was further evaluated by phylogenetic analysis. Nucleotide sequences of *V. modesta cyclotide* genes were used as a query in BLASTn searches against GenBank database for possible homologues in other *Viola* species. A total of 30 *Viola cyclotide* genes from three Viola species in NCBI database in addition to 14 genes from the current study were used for phylogenetic analysis.

The phylogenetic tree indicated that *Violaceae cyclotide* genes were divided into two main subfamilies: the Mobius and bracelet (Figure 4). Most of cyclotides belong to bracelet subfamily. In this work, we designed two primers based on sequences from both subfamilies. However, all *cyclotide* genes obtained in this work belonged to bracelet subfamily. *Cyclotide* genes in bracelet subfamily distributed in six clades. *V. modesta cyclotide* genes were classified in two clades. *Vmcyc5*、*Vmcyc10*, *Vmcyc70*, *Vmcyc7*, *Vmcyc40* and *Vmcyc11* in one clade separate from other cyclotides. *Vmcyc9*, *Vmcyc3*, *Vmcyc16*, *Vmcyc17*, *Vmcyc4*, *Vmcyc12*, and *Vmcyc8* were in another clade with *Voc.1* (*V. odorata* cyclotide 1), *Voc.2* (*V. odorata* cyclotide 2), Voc.3 (*V. odorata* cyclotide3), Vbc.3d (*V. baoshanensis* cyclotide 3d). The first group consisted of 98-100 residues and the ER-region contained a conserved sequence consisting of ATAFALP, while sequences in the other clade consisted of 104 residues and a conserved sequence consisting of AAFALP.

Cyclotide genes from different *Viola* species separated in one or two clades. These data suggested that the *Vmcyc* genes in *Viola* were ancient with multi-origins.

Expression Pattern of *Vmcyc1* and *Vmcyc7* in Different Tissues

Since *V. modesta* cyclotides were classified in two clades and cyclotides within each clade were very similar, we chose one cyclotide in each clade: *Vmcyc1* and *Vmcyc7*. Specific primers were designed for each cyclotide. Since we did not have the sequence of any control gene from *V. modesta*, tubulin sequence from *Ricinus communis* ,*Populus trichocarpa, V. cornuta, Lotus japonicas* from the family *Fabaceae* were aligned and two primers were designed. A segment of 450 bp was amplified and cloned and specific primers were designed. The isolated tubulin fragment showed the highest similarity with its orthologs gene in *V. cornuta*. Expression levels of the *Vmcyc1* and *vmcyc7* genes were examined in different tissues including roots, stems, leaves, flowers, seeds, peduncles, and capsules using quantitative reverse transcription-PCR. Normalization was carried out based on tubulin expression as housekeeping gene. *Vmcyc1* expression in tissues was different (Figure 5). The highest transcript level was observed in peduncles and capsules, while the lowest transcript level was detected in leaves and flowers. The transcript level of *Vmcyc7* was also differing in various tissues: the highest level was observed in flowers and capsules and the lowest in stems (Figure 6).

DISCUSSION

Cyclotides sequences have been isolated based on different methods including extraction of cyclotides protein using MS.MS, 2D NMR and screening a cDNA library or RT-PCR. Cyclotide precursor includes parts that is conserved and can be used for primer design. A degenerative forward primer encoding a conserved region (AAFALPA) of ER-signals in the *Violaceae* cyclotide proteins have been used as a powerful strategy for cloning of nearly full length precursors of *cyclotide* genes (Mulvenna *et al*., 2005; Simonsen *et al*., 2005). By screening a cDNA library of *V. baoshanensis* roots and using RACE and RT-PCR methods, 23 cDNA clones were identified (Zhang *et al*., 2009). Twelve cyclotides from two Panamanian species,

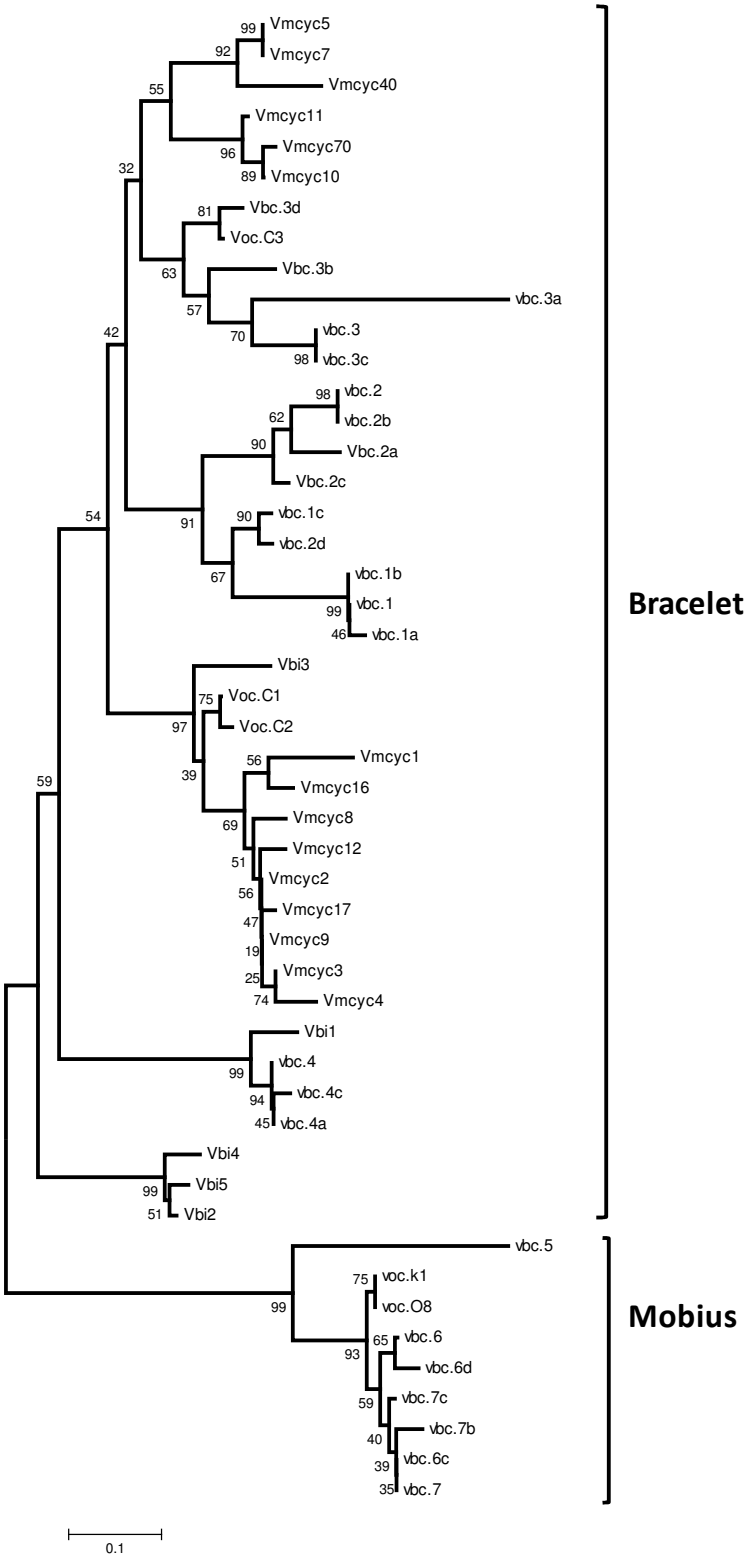

Figure 4. Phylogenetic tree of cyclotide precursors found in the *Violaceae* plant family. The bootstrap consensus tree was made by Mega4.0. Numbers next to the nodes give bootstrap values. The tree comprises all already reported cyclotide precursor sequences, and *V. modesta* cyclotides (Table 1).

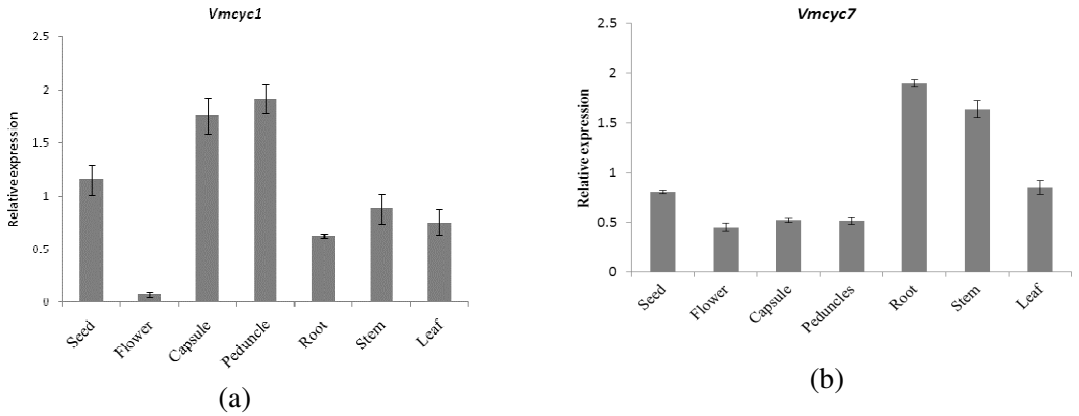

Figure 5. Relative RT-PCR of *Vmcyc1* (a) Relative RT-PCR of *Vmcyc7*(b)in *V. modesta* organs. mRNA were determined relative to that of *Tubolin* gene. Mean values are shown with standard error bars which were calculated based on three replications.

Gloeospermum pauciflorum Hekking and *Gloeospermum blakeanum* (Standl.) were characterized through cDNA screening (Burman *et al*., 2010). The sequences of 11 cyclotides, were determined by isolation and MS.MS sequencing of proteins and screening of a cDNA library of *V. biflora* in parallel (Herrmann *et al*., 2007). They used a degenerate primer against a conserved (AAFALPA) motif in the cyclotide precursor ER signal sequence, which yielded a series of predicted cyclotide sequences that were correlated to those of the isolated proteins. In this study, 14 *cyclotide* genes were identified by screening a *V. modesta* cDNA using 3'RACE. After analyzing the ER signal regions of these *cyclotide* genes, two conserved sequences were found, including ATFALPS (F) and AAFALPA with the last one incorporating a known conserved-region (AAFALPA). Conserved region ATFALPS (F) is most likely new and has not been reported yet in other *Viola* species.

Our first primer was based on conserved region (AAFALPA) which was found in both bracelet and Mobius precursors. All of cyclotide precursors of *V. modesta* belonged to bracelet subfamily. We used another primers based on AAFALPA-AFA conserved motif, which can amplify Mobius subfamily (Zhang *et al*., 2009), but still got the bracelet sequences. Other works also have shown that utilizing the AAFALPA primer (Simonsen *et al*., 2005; Herrmann *et al*., 2008) yields bracelet precursors. Of the 150 sequences found in the Cybase server (http://www.cybase.org.au), the majority of cyclotides (> 67%) belong to the bracelet subfamily (Simonsen *et al*., 2005). Most of cloned cyclotide using RACE and RT-PCR belonged to bracelet subfamily (Trabi and Criak 2004; Tang *et al*., 2010). No detection of Mobius cyclotide in *V. modesta* may be related to low expression of *Mobius* genes in tissues.

In this study, 14 bracelet cyclotides sequences were predicted from *V. modesta*. By constructing a phylogenetic tree combining cyclotides from *Violaceae* in NCBI database and the above predicted cyclotides, it is clear that these *Violaceae* cyclotides are classified in two clear classes, namely, bracelet and Mobius, consistent with previous observations that in general bracelet cyclotides are more common than Mobius cyclotides. Bracelets were further separated to six subclasses and *V. modesta cyclotide* genes were classified in two clades. The bracelets are the more structurally diverse compared to Mobius. In addition, the bracelet cyclotides are generally more cytotoxic than Möbius cyclotides (Lindholm *et al*., 2002; Svangard *et al*., 2004; Herrmann *et al*., 2008). Within the bracelet subfamily, it can be noted that

cyclotides with several positive residues in loops 5 and 6 have higher activity than those lacking such residues (Burman *et al*., 2010).

Recent studies have shown that cyclotide expression varies in different tissues. The expression patterns analysis of cyclotides in various *Viola* species (*Violaceae*) showed tissue specificity (Trabi *et al*., 2004). For example, they have isolated a cyclotide which is only expressed in underground parts of *V. hederaceae*. Tissue dependent expression of *cyclotide* genes has been observed in VbCP1S-7S of *V. baoshanensis,* which indicates that cyclotide expression could be regulated at transcript level (Zhang *et al*., 2009). A leaf-specific cyclotide (vhl-1) isolated from *V. hederaceae* expressed 31-residue cyclotide (Chen *et al*., 2005). The high expression levels of cyclotide precursor transcripts were detected in *Oldenlandia affinis* leaf transcriptome (Qin *et al*., 2010). In total, 31 ESTs encoded cyclotide precursors, representing a distinct commitment of 2.8% of the transcriptome to cyclotide biosynthesis with the abundance of mature cyclic peptides in *O. affinis*. *V. modesta* cyclotides also showed tissue specific pattern. In addition, *Vmcyc*1 and *Vmcyc2,* which belong to two separate subclasses, were differentially expressed in different tissues.

Most of cyclotides discovered to date have shown wide bioactivities including insecticidal (Jennings *et al*., 2001), antimicrobial activities (Tam *et al*., 1999), anti-HIV (Chen *et al*., 2005), cytotoxic (Lindholm *et al*., 2002), hemolytic (Chen *et al*., 2006), neurotensin antagonism (Witherup *et al*., 1994), uterotonic (Gran, 1973) and trypsin inhibitor (Hernandez *et al*., 2000). Also, the cyclotides are thought to be a valuable peptide-based template for drug design and agrochemical applications (Craik *et al*., 1999). *V. modesta* cyclotides showed significant homology to other known *Violaceae* cyclotides. More detailed analysis of these genes and finding promising candidates for genetic engineering purpose is needed. In addition, we have enriched the knowledge of the cyclotide family by characterizing novel cyclotides from this family.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Faculty of Agriculture, University of Kurdistan, for providing the funds and research facilities.

REFERENCES

1. Bendtsen, J. D., Nielsen, H., Von Heijne, G. and Brunak, S. 2004. Improved Prediction of Signal Peptides. SignalP 3.0. *J. Mol. Biol.,* **340:** 783-795.
2. Burman, R., Gruber, C. W., Rizzardi, K., Herrmann, A., Craik, D. J., Gupta, M. P. and Goransson, U. 2010. Cyclotide Proteins and Precursors from the Genus *Gloeospermum*: Filling a Blank Spot in the Cyclotide Map of *Violaceae*. *Phytochem.,* **71:** 13-20
3. Chen, B., Colgrave, M. L., Daly, N. L., Rosengren, K. J., Gustafson, K. R. and Craik, D. J. 2005. Isolation and Characterization of Novel Cyclotides from *Viola hederaceae*: Solution Structure and Anti-HIV Activity of vhl-1, a Leaf-specific Expressed Cyclotide. *J. Biol. Chem.,* **280:** 22395-22405
4. Craik, D. J. 2012. Host-defense Activities of Cyclotides. *Toxin.,* **4:** 139-156
5. Craik, D. J., Daly, N. L. Bond, T. and Waine, C. 1999. Plant Cyclotides: A Unique Family of Cyclic and Knotted Proteins that Defines the Cyclic Cystine Knot Structural Motif. *J. Mol. Biol.,* **294:** 1327-1336.
6. Craik, D. J., Daly, N. L. and Waine, C. 2001. The Cystine Knot Motif in Toxins and Implications for Drug Design. *Toxicon.,* **39:** 43-60.
7. Daly, N. L., Love, S., Alewood, P. F. and Craik, D. J. 1999. Chemical Synthesis and Folding Pathways of Large Cyclic Polypeptides: Studies of the Cystine Knot Polypeptide Kalata B1. *Biochem.,* **38:** 10606-10614.
8. Dutton, J. L., Renda, R. F. Waine, C., Clark, R. J., Daly, N. L., Jennings, C. V., Anderson, M. A. and Craik, D. J. 2004. Conserved Structural and Sequence Elements Implicated in the Processing of Gene-encoded Circular Proteins. *J. Biol. Chem.,* **279:** 46858-46867.

9. Erben, M. 1996. The Significance of Hybridization on the Forming of Species in the Genus *Viola*. *Bocconea.,* **5:** 113-118.
10. Gerlach S. L. and Mondal, D. 2012. The Bountiful Biological Activities of Cyclotides. *Chron. Young Scientist,* **3:** 169.
11. Gran, L. 1973. Oxytocic Principles of *Oldenlandia affinis*. *Lloydia* **36:** 174-178.
12. Herrmann A., E. Svangard, J. Gullbo, L. Bohlin, and U. Goransson. 2006. Key Role of Glutamic Acid for the Cytotoxic Activity of the Cyclotide Cycloviolacin O2. *Cell. Mol. Life. Sci.,* **63:** 235-245.
13. Herrmann. A., Burman, R., Mylne, J. S., Karlsson, G., Gullbo, J., Craik, D. J., Clark, R. J. and Goransson, U. 2008. The Alpine Violet (*Viola biflora*), Is a Rich Source of Cyclotides with Potent Cytotoxicity. *Phytochem.,* **69:** 939-952.
14. Hernandez, J. F., Gagnon, J., Chiche, L., Nguyen, T. M., Andrieu, J. P., Heitz, A., Hong, T. T., Pham, T. T. C., and Nguyen, D. L. 2000. Squash Trypsin Inhibitors from *Momordica cochinchinensis* Exhibit an Atypical Macrocyclic Structure. *Biochem.,* **39:** 5722–5730.
15. Jennings, C., West, J., Waine, C., Craik, D. and Anderson, M. 2001. Biosynthesis and Insecticidal Properties of Plant Cyclotides: The Cyclic Knotted Proteins from *Oldenlandia affinis*. *Proc. Natl. Acad. Sci. USA,* **98:** 10614–10619.
16. Lindholm, P., Goransson, U., Johansson, S., Claeson, P., Gullbo, J., Larsson, R., Bohlin, L. and Backlund, A. 2002. Cyclotides: A Novel Type of Cytotoxic Agents. *Mol. Cancer. Ther.,* **1:** 365-369.
17. Mazzara, M. and James, D. J. 2000. The Influence of Photoperiodic Growth Condition on Isolation of RNA from Strawberry (*Fragaria×ananassa* Duch.) Tissue. *Mol. Biotech.,* **15:** 237-241.
18. Mulvenna, J. P., Sando, L. and Craik, D. J. 2005. Processing of a 22 kDa Precursor Protein to Produce the Circular Protein Tricyclon A. *Structure,* **13:** 691-701.
19. Qin, Q., McCallum, E. J., Kaas, Q., Suda, J., Saska, I., Craik, D. J. and Mylne, J. S. 2010. Identification of Candidates for Cyclotide Biosynthesis and Cyclisation by Expressed Sequence Tag Analysis of *Oldenlandia affinis*. *BMC Genomic.,* **11:** 111.
20. Simonsen, S. M., Sando, L., Ireland, D. C., Colgrave, M. L., Bharathi, R., Goransson, U. and Craik, D. J. 2005. A Continent of Plant Defense Peptide Diversity: Cyclotides in Australian *Hybanthus* (*Violaceae*). *Plant Cell.,* **17:** 3176-3189.
21. Svangard, E., Goransson, U., Hocaoglu, Z., Gullbo, J., Larsson, R., Claeson, P. and Bohlin, L. 2004. Cytotoxic Cyclotides from *Viola tricolor*. *J. Nat. Prod.,* **67:** 144-147.
22. Tam, J. P., Lu, Y. A., Yang, J. L. and Chiu, K. W. 1999. An Unusual Structural Motif of Antimicrobial Peptides Containing End-to-end Macrocycle and Cystine-knot Disulfides. *Proc. Natl. Acad. Sci. USA,* **96:** 8913–8918.
23. Tang, J., Wang, C. K., Pan, X., Yan, H., Zeng, G., Xu, W., He, W., Daly, N. L., Craik, D. J. and Tan, N. 2010. Isolation and Characterization of Cytotoxic Cyclotides from *Viola tricolor*. *Peptide.,* **31:** 1434-1440.
24. Trabi, M. and Craik, D. J. 2004. Tissue-specific Expression of Head-to-tail Cyclized Miniproteins in *Violaceae* and Structure Determination of the Root Cyclotide *Viola hederacea* Root Cyclotide 1. *Plant Cell.,* **16:** 2204-2216.
25. Trabi, M., Svangard, E., Herrmann, A., Göransson, U., Claeson, P., Craik, D. J. and Bohlin, L. 2004. Variations in Cyclotide Expression in *Viola* Species. *J. Nat. Prod.,* 67:806-810.
26. Wang, C. K. L., Kaas, Q., Chiche, L. and Craik, D. J. 2008. CyBase: A Database of Cyclic Protein Sequences and Structures, with Applications in Protein Discovery and Engineering. *Nucleic Acids. Res.,* **36:** 206-210.
27. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. 1994. Cyclopsychotride A, a Biologically Active, 31-residue Cyclic Peptide Isolated from *Psychotria longipes*. *J. Nat. Prod.* **57:** 1619-1625.
28. Zhang, J., Liao, B., Craik, D. J., Li, J. T., Hu, M. and Shu, W. S. 2009. Identification of Two Suites of *Cyclotide Precursor* Genes from Metallophyte *Viola baoshanensis*: cDNA Sequence Variation, Alternative RNA Splicing and Potential Cyclotide Diversity. *Gene,* **431:** 23-32.

همسانه سازی و آنالیز ژنهای سایکلوتاید در گیاه *Viola modesta* Fenzl

ن. کداری، ب. بهرام نژاد، ج. رستم زاده، ه. معروفی، و س. ترکمان

چکیده

سایکلوتایدها دسته ای از پپتیدهای حلقوی هستند که غنی از باندهای دی سلفیدی می باشند. این پروتیین ها فعالیت‌های زیستی متنوع دارند و در گیاهان مناطق گرم در خانواده های ,Rubiaceae Violaceae, Cucurbitaceae and Fabaceaeگزارش شده اند . در این پژوهش ژن‌های کد کننده پروتئین‌های سایکلوتاید در *Viola modesta* گونه بومی آسیای غربی در استان کردستان ایران همسانه سازی و تعیین توالی شدند. در نهایت پس از تعیین توالی، ۱۴ توالی‌های نوکلئوتیدی بدست آمد که با توالی موجود در پایگاه دادهNCBI همتراز و مشخص شد توالی‌های بدست آمده سایکلوتاید بودند .سایکلوتایدهای بدست آمده در گروه بریسلت قرار گرفتند .سایکلوتایدهای گیاه *Viola modesta* بر اساس رسم دندروگرام و مقایسه با سایر سایکلوتایدها، در گروه بریسلت به دو دسته تقسیم شدند. بیان دو سایکلوتاید Vmcyc1 و Vmcyc7 به عنوان نماینده از ۱۴ سایکلوتاید بدست آمده از*Viola modesta* در اندام‌های گل، دم‌گل، ساقه، ریشه، کپسول و برگ بررسی شد و مشخص شد کهVmcyc7در همه اندام‌ها بیان شد و بیان آندر کپسول و گل بیشتر از سایر قسمت‌ها بود و Vmcyc1 در کپسول و محور گل بیشترین میزان بیان و در گل کمترین میزان بیان را داشت.