Journal of Plant Physiology 178 (2015) 17–26

Contents lists available at ScienceDirect

Journal of Plant Physiology

journal homepage: www.elsevier.com/locate/jplph

Molecular biology

Transcriptomic screening for cyclotides and other cysteine-rich proteins in the metallophyte *Viola baoshanensis*

Jun Zhang[a,b], Jintian Li[b], Zebo Huang[a], Bing Yang[b], Xiaojie Zhang[a], Dehua Li[a], David J. Craik[c], Alan J.M. Baker[d], Wensheng Shu[b], Bin Liao[b,*]

[a] *Guangdong Pharmaceutical University, School of Biosciences and Biopharmaceutics, Guangdong Province Key Laboratory for Biotechnology Drug Candidates, Guangzhou 510006, China*
[b] *Sun Yat-sen University, School of Life Sciences, State Key Laboratory of Biocontrol, Guangzhou 510006, China*
[c] *The University of Queensland, Institute for Molecular Bioscience, Brisbane 4072, QLD, Australia*
[d] *The University of Melbourne, School of Botany, Parkville 3010, VIC, Australia*

ARTICLE INFO

Article history:
Received 19 August 2014
Received in revised form 1 January 2015
Accepted 28 January 2015
Available online 26 February 2015

Keywords:
Cyclotides
Cysteine-rich proteins
CrpExcel
Transcriptome profiles
Viola baoshanensis

ABSTRACT

Cysteine (Cys)-rich proteins (CRPs) are frequently associated with plant defense and stress resistance. *Viola baoshanensis* is a cadmium (Cd) hyper-accumulating plant whose CRPs-based defense systems are so far poorly understood. Next generation sequencing (NGS) techniques and a specialist searching tool, CrpExcel, were employed for identifying CRPs in *V. baoshanensis*. The transcriptome sequences of *V. baoshanensis* were assembled primarily from 454FLX/Hiseq2000 reads of plant cDNA sequencing libraries. CrpExcel was then used to search the ORFs and 9687 CRPs were identified, and included zinc finger (ZF) proteins, lipid transfer proteins, thaumatins and cyclotide precursors. Real-time PCR results showed that all CRP genes tested are constitutively expressed, but the genes of defensive peptides showed greater up-regulated expression than those of ZF-proteins in Cd- and/or wounding (Wd) treatments of *V. baoshanensis* seedlings. The NGS-derived sequences of cyclotide precursor genes were verified by RT-PCR and ABI3730 sequencing studies, and 32 novel cyclotides were identified in *V. baoshanensis*. In general, the metal-binding sites of ZF-containing CRPs also represented the potential vulnerable targets of toxic metals. This study provides broad insights into CRPs-based defense systems and stress-vulnerable targets in *V. baoshanensis*. It now brings the number of cyclotide sequences in *V. baoshanensis* to 53 and based on projections from this work, the number of cyclotides in the Violaceae is now conservatively estimated to be >30000.

Introduction

Plants have evolved a series of defense strategies for adapting to various soil conditions, pathogen infection or herbivore attack (Zhang et al., 2009a; Acharya et al., 2013). For example, cysteine (Cys) is one of the products of plant sulfur metabolism, and one of its functions, via the formation of Cys-rich proteins (CRPs), is in plant stress resistance (Mirouze et al., 2006; Kunihiro et al., 2013). CRPs are widely expressed in plants as part of their defense arsenals and signal regulators, and include zinc finger (ZF) proteins, metallothioneins (MTs) and plant defense peptides (Terras et al., 1995; Simonsen et al., 2005). ZF-proteins typically possess a metal-binding site and function in plant regulatory networks associated with stress resistance (Ciftci-Yilmaz and Mittler, 2008). The MTs are the most well-documented metal-binding proteins for heavy metal homeostasis in plants (Cobbett and Goldsbrough, 2002). In contrast, cyclotides, defensins, thionins, and snakins are common Cys-rich peptides involved in defense against toxic heavy metals and/or a number of biotic stresses (Abdallah et al., 2010; Ng et al., 2012). Thus, screening for CRPs is important for exploring plant defense systems.

Cyclotides are head-to-tail cyclic peptides containing 28–37 amino acids (Fig. 1A and B) (Craik et al., 1999; Daly et al., 2009). Over 280 wild-type cyclotides are documented in CyBase (Mulvenna et al., 2006; Wang et al., 2008), a web-accessible

---

Abbreviations: Cd, cadmium; CRPs, Cysteine-rich proteins; Cys, Cysteine; ER, endoplasmic reticulum; MTs, Metallothioneins; NCBI, National Center for Biotechnology Information; NGS, next generation sequencing; NTR, N-terminal repeat; qRT-PCR, quantitative real-time PCR; Wd, Wounding; ZF, zinc finger.

* Corresponding author.
E-mail addresses: zhangj80@aliyun.com (J. Zhang), lijtian@mail.sysu.edu.cn (J. Li), zebohuang@gdpu.edu.cn (Z. Huang), andyang@tom.com (B. Yang), zhangxiaojie9205@163.com (X. Zhang), dehuali_25@163.com (D. Li), d.craik@imb.uq.edu.au (D.J. Craik), ajmb@unimelb.edu.au (A.J.M. Baker), shuws@mail.sysu.edu.cn (W. Shu), lsslb@mail.sysu.edu.cn (B. Liao).

http://dx.doi.org/10.1016/j.jplph.2015.01.017
0176-1617/

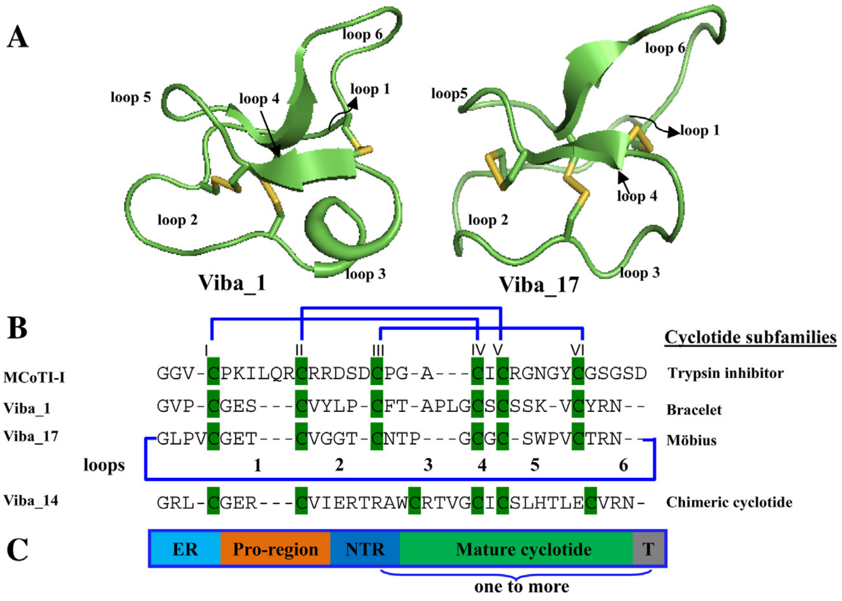

Fig. 1. Structural frameworks, sequences and precursors of cyclotides. The four representative cyclotides shown are in CyBase (Wang et al., 2008), and among them, Viba_1, 14 and 17 were identified from *V. baoshanensis* in our previous study (Zhang et al., 2009b). (A) Backbone ribbon and the disulfide network of Viba_1 and Viba_17. The disulfide bonds are shown with yellow lines and the loops are labeled 1–6. (B) Alignment of representative cyclotides from the trypsin inhibitor, bracelet and Möbius subfamilies, as well as a chimeric cyclotides. Their disulfide bonds are shown by blue lines connecting the Cys residues. (C) Block diagram of cyclotide precursor proteins. The precursors include an endoplasmic reticulum (ER) signal region and a pro-region, followed by one or more domains comprising an N-terminal repeat (NTR), a mature cyclotide and a short C-terminal tail (T).

database on circular proteins. Cyclotides are biosynthesized from the post-translational processing of gene-encoding precursor proteins containing one or more mature cyclotide domains (Fig. 1C) (Jennings et al., 2001; Zhang et al., 2009b). They appear to be widely distributed in flowering plants (Gruber et al., 2008; Gerlach et al., 2013; Koehbach et al., 2013a), having been reported so far in the plant families Violaceae, Rubiaceae, Fabaceae, Cucurbitaceae, Solanaceae and Poaceae (Plan et al., 2007; Zhang et al., 2014), albeit with a selective distribution in these families (Trabi and Craik, 2004; Avrutina et al., 2005; Mulvenna et al., 2006).

Cyclotides display metal-binding (Shenkarev et al., 2008; Wang et al., 2009) and immuno-suppressive activities (Gründemann et al., 2012, 2013), and also exhibit a diverse range of bioactivities related closely to plant defense, including anti-viral (Bokesch et al., 2001; Daly et al., 2004), cytotoxic (Lindholm et al., 2002), antibacterial (Tam et al., 1999) and anti-insect (Jennings et al., 2005; Gruber et al., 2007; Barbeta et al., 2008). Owing to their rich sequence variability, stable structures and varied functions, cyclotides are regarded as important defense peptides in plants (Mylne et al., 2010). Furthermore, cyclotides have been produced by phytochemical extraction, chemical synthesis, and/or recombinant expressions for developing drugs and agrichemical products (Craik et al., 1999; Koehbach et al., 2013b; Poth et al., 2013).

Viola baoshanensis Shu, Liu et Lan is a newly identified species of the genus *Viola* (Shu et al., 2003) that has been identified not only as a cadmium (Cd)-hyper-accumulator, but also a strong Pb/Zn accumulator based on field surveys, pot experiments and hydroponic culture (Liu et al., 2004; Wu et al., 2010). *V. baoshanensis* is therefore an ideal model for exploring the molecular physiology of plant defense and heavy metal resistance (Li et al., 2010). In a previous study of this plant, several cyclotide genes were identified using a root subtractive cDNA library, including six CRP genes (Zhang et al., 2009a). In total, 21 cyclotide sequences were identified by RACE and RT-PCR techniques (Zhang et al., 2009b). However, it is still unclear how many more cyclotides and other CRP types are expressed in *V. baoshanensis*.

In recent years next-generational sequencing (NGS) platforms, such as Roche 454FLX sequencers and Illumina Hiseq sequencers, have been widely used for the high-throughput sequencing of many organisms (Yu et al., 2012; Gao et al., 2013). 454FLX and Hiseq techniques are often used together for genomic and/or transcriptome sequencing of non-model plants because of their different sequencing length and depth (Gao et al., 2013). Theoretically, the two NGS techniques can facilitate plant CRP research by producing a great number of nucleotide sequences from plant sequencing libraries. However, this is still a challenging task because of the lack of user-friendly methods for extracting known or unknown CRPs from a large number of nucleotide/protein sequences. In the present study, transcriptomes of *V. baoshanensis* were sequenced using Roche 454FLX and Illumina Hiseq2000 sequencing platforms and a Microsoft Excel-based tool, CrpExcel (Supplement 1), was developed for searching plant-level CRPs from the ORFs of the NGS-derived transcriptome sequences of *V. baoshanensis*.

Materials and methods

Plant culture and RNA extraction

A rapid micro-propagation system has been developed for *Viola baoshanensis* (Li et al., 2010), and seedlings produced from tissue cultures were used in this study. Seedlings grown to shoot lengths of 3–4 cm were selected for hydroponic tests in half-strength Hoagland's solution (pH 5.8–6.0). The seedlings were cultured at 25 °C with 16 h light illumination in a glasshouse, and the culture solutions replaced at 2-day intervals. To explore the transcriptome information associated with Cysteine (Cys)-rich proteins (CRPs) of *V. baoshanensis*, the seedlings were treated by both shearing each leaf and/or adding 300 μM $CdCl_2$ in the hydroponic solution for 2 days, as described in previous studies (Zhang et al., 2009a; Wu et al., 2010). After treatment, the leaves were harvested for RNA extraction and next generation sequencing (NGS) analysis.

To evaluate the expression patterns of the CRP genes of interest, seedlings were cultured in solutions: (1) without treatment (Control group, CK); (2) supplemented with 300 μM $CdCl_2$ (cadmium (Cd)-treated group); (3) wounded by shearing each