Genomics Proteomics Bioinformatics 11 (2013) 172–181

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

ORIGINAL RESEARCH

Identification of ta-siRNAs and *Cis*-nat-siRNAs in Cassava and Their Roles in Response to Cassava Bacterial Blight

Andrés Quintero, Alvaro L. Pérez-Quintero, Camilo López *

Departamento de Biología, Universidad Nacional de Colombia, Bogotá DC 111321, Colombia

Received 8 October 2012; revised 20 February 2013; accepted 14 March 2013
Available online 7 May 2013

KEYWORDS

ta-siRNA;
nat-siRNA;
Natural antisense transcript;
Cassava;
Xanthomonas

Abstract *Trans*-acting small interfering RNAs (ta-siRNAs) and natural *cis*-antisense siRNAs (*cis*-nat-siRNAs) are recently discovered small RNAs (sRNAs) involved in post-transcriptional gene silencing. ta-siRNAs are transcribed from genomic loci and require processing by microRNAs (miRNAs). *cis*-nat-siRNAs are derived from antisense RNAs produced by the simultaneous transcription of overlapping antisense genes. Their roles in many plant processes, including pathogen response, are mostly unknown. In this work, we employed a bioinformatic approach to identify ta-siRNAs and *cis*-nat-siRNAs in cassava from two sRNA libraries, one constructed from healthy cassava plants and one from plants inoculated with the bacterium *Xanthomonas axonopodis* pv. manihotis (*Xam*). A total of 54 possible ta-siRNA loci were identified in cassava, including a homolog of *TAS3*, the best studied plant ta-siRNA. Fifteen of these loci were induced, while 39 were repressed in response to *Xam* infection. In addition, 15 possible *cis*-natural antisense transcript (*cis*-NAT) loci producing siRNAs were identified from overlapping antisense regions in the genome, and were found to be differentially expressed upon *Xam* infection. Roles of sRNAs were predicted by sequence complementarity and our results showed that many sRNAs identified in this work might be directed against various transcription factors. This work represents a significant step toward understanding the roles of sRNAs in the immune response of cassava.

Introduction

Small RNAs (sRNAs) have emerged as important factors in the regulation of gene expression, which are involved in transcriptional gene silencing (TGS) and post-transcriptional gene silencing (PTGS) through inactivation, degradation or translational repression of mRNAs [1,2]. One of the best known types of plant sRNAs is microRNAs (miRNAs). miRNAs originate from nuclear genomic loci and are transcribed by RNA polymerase II as primary miRNAs (pri-miRNAs), which are processed by Dicer like-1 (DCL1) to form mature miRNAs

* Corresponding author.
E-mail: celopezc@unal.edu.co (López C).

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.gpb.2013.03.001

(21 nt) [1,2]. The mature miRNAs are then incorporated into an RNA-induced silencing complex (RISC), where they guide the degradation of complementary mRNAs [1,2].

Alternatively, miRNAs can target other non-coding RNAs (ncRNAs) to generate a type of plant-specific small interfering RNAs called *trans*-acting small interfering RNAs (ta-siRNAs) [3,4]. ta-siRNA loci (*TAS*) are transcribed from nuclear genes and then recognized and cleaved by Argonaute (AGO)-containing RNA-induced silencing complex (RISC) coupled with particular miRNAs [3,4]. The cleaved products are then used by an RNA-dependent RNA polymerase (RDR6) to produce double-stranded RNAs (dsRNAs), which are sequentially cleaved by DCL proteins to produce phased, ~20–24 nt ta-siRNAs [3,4]. These ta-siRNAs function similarly to miRNAs by targeting complementary genes and initiating their degradation or translation arrest [3,4]. Each *TAS* is cleaved by one or two miRNAs, producing several ta-siRNAs in phase and resulting in an amplified signal that silences several target molecules [3,4]. To date, 8 *TAS* loci have been identified and validated in *Arabidopsis*. These loci are grouped into 4 families, *TAS1–4*, each with a distinct biogenesis and function. The *TAS1* family is comprised of 3 highly similar genes that target 5 transcripts with unknown function [3,4]. The *TAS2* family contains ta-siRNAs that target mRNAs coding for pentatricopeptide repeat (PPR) proteins [4]. The cleavage of both *TAS1* and *TAS2* transcripts is directed by miR-173 in conjunction with AGO1 [3,5]. *TAS3* is highly conserved in plants, including monocots and dicots, and cleaved through the combined action of miR-390 and AGO7 [6–8], producing ta-siRNAs that target auxin-responsive factors (ARF) [6–8]. *TAS3*, its targets (ARF2–4) and miR-390 are involved in a self-regulatory loop, controlling lateral root growth [9,10]. In this loop, the ta-siRNAs produced from *TAS3* inhibit the expression of ARF2/3/4, ARF4 prevents miR-390 accumulation [9,10], while ARF3 up-regulates miR-390 accumulation in response to auxin [9,10]. The cleavage of *TAS*4 is triggered by miR-828, producing a series of ta-siRNAs that target several MYB transcription factors including PAP1, PAP2 and MYB113 [11,12].

In addition to miRNAs and ta-siRNAs, there are other gene-silencing mechanisms involving sRNAs, such as nat-siRNAs from natural antisense transcripts (NATs), which are pairs of endogenous coding or non-coding RNAs with perfect complementarity to each other [13]. There are two classes of nat-siRNAs: *cis* and *trans*. The former are derived from NATs located in the opposite strand of the same genomic locus. They originate when both transcripts in the NAT-pair are simultaneously transcribed, and their complementary regions match up forming a dsRNA, which is employed as a template to produce siRNAs [14]. The siRNA processing is dependent on DCL1 and/or DCL3, the siRNAs are predominantly produced from only one of the transcripts and they direct the cleavage of the transcript produced in the opposite strand. In this manner the transcripts involved in the *cis*-NAT pair are reciprocally regulated [15]. On the other hand, *trans*-nat-siRNAs arise from two overlapping transcripts that are located in different loci, and are processed in a similar manner as *cis*-nat-siRNAs [13].

The functions of ta-siRNA and *cis*-nat-siRNAs are largely unknown. However, some aspects of their overall importance in various plant processes have been established. A crucial role for ta-siRNAs has been reported for example in pollen maturation [16], in response to low temperature [17] and hypoxia [18]. Likewise, *cis*-nat-siRNAs play an important role in regulating the expression of the genes from which they arise; that is, if one of the genes is induced, the other is repressed [14]. This phenomenon has been observed in a wide variety of gene pairs with various functions [14]. For example, it has been shown in *Arabidopsis* that a sperm-specific *cis*-NAT pair, formed by KPL and its inversely transcribed gene ARIADNE14, is regulated by nat-siRNAs, which is required for the fertilization process. In the absence of KPL, no nat-siRNAs are produced, resulting in the induced regulation of ARIADNE14 RNA, increasing its levels and in consequence inhibiting fertilization [19].

While, some plant miRNAs have been found to have a direct role in defense against bacteria, fewer studies have focused on the role of ta-siRNAs or *cis*-nat-siRNAs. It is known, for example, that upon microbe-associated molecular pattern (MAMP)-triggered immunity (MTI), miR-393 is induced after recognition of the MAMP flagelin by the FLS2 receptor, which consequently targets ARFs and disrupts auxin signaling [20]. Additionally, other miRNA families also contribute to antibacterial defense by modulating hormonal responses [21,22].

There are, however, some indications that ta-siRNAs and nat-siRNAs may also be important for antibacterial defense. It was shown that miRNAs could target the tobacco *N* resistance gene and produce ta-siRNAs [23]. In addition, several ta-siRNAs target and regulate the expression of multiple NB-LRR-coding resistance genes in legumes during their interaction with *Rhizobium* [24]. On the other hand, siRNAs derived from a *cis*-NAT locus are induced in *Arabidopsis* in response to the effector AvrRpt2 from *Pseudomonas syringae* pv. tomato DC3000 [25].

Cassava (*Manihot esculenta*) is a staple crop that represents the main source of carbohydrates for more than 600 million people worldwide. Cassava starch also has a high potential for bioethanol production. Cassava is grown mainly by farmers in marginal areas of South America, Africa and Asia, where its production is severely affected by the bacterium *Xanthomonas axonopodis* pv. manihotis (*Xam*) [26]. Previously, we reported the identification of cassava miRNAs that were induced or repressed after inoculation with *Xam* through the sequencing of sRNA libraries [27]. Here, we expand our study on bacterial response of cassava by examining the function of small ncRNAs through identification of differentially expressed ta-siRNA and *cis*-nat-siRNAs.

Results

Identification of conserved ta-siRNAs in cassava

With the aims of establishing a set of parameters suitable for ta-siRNA prediction, the expression profiles of known *Arabidopsis TAS* loci and their conserved counterparts in cassava were analyzed.

We used the reported *Arabidopsis TAS* sequences as a query to screen the cassava genome for loci that have a similar sequence and are also predicted to have miRNA cleavage sites. Only one locus showed high similarity to an *Arabidopsis* locus: *TAS3*. The overall nucleotide identity was 54% over the length of the *Arabidopsis* locus and the location of this conserved *TAS* in the latest version of the cassava genome was scaffold00708:185488,185707. From now on, this sequence is