Rapid Communication

Characterization of the siRNAs associated with peach latent mosaic viroid infection

Patrick St-Pierre a,1, I. Fekih Hassena,1, D. Thompson b, J.P. Perreault a,⁎

a RNA group/Coupe an, Département de biochimie, Université de Sherbrooke, Sherbrooke, Québec, Canada J1H 5N4
b Center for Plant Health, Canadian Food Inspection Agency, Sidney, British Columbia, Canada V8L 1H3

ARTICLE INFO

Article history:
Received 12 September 2008
Returned to author for revision 25 October 2008
Accepted 8 November 2008
Available online 9 December 2008

Keywords:
Viroid
RNA silencing
Post-transcriptional gene silencing
Sequencing
Small non-coding RNA

ABSTRACT

The peach latent mosaic viroid (PLMVd) is a small, circular RNA species that has been shown to induce RNA silencing in plants. With the goal of better understanding the biological mechanism underlying this process, the siRNAs found in PLMVd infected peach leaves were isolated and sequenced. Analysis of the resulting data prompted several conclusions, including: i. PLMVd strands of both polarities are substrates for the Dicer-like enzymes found in peach leaves; ii. the more highly structured regions of PLMVd trigger the activity of the Dicer-like enzymes; and, iii. the circular PLMVd conformers appear to be favored for transport into the cytoplasm.

© 2008 Elsevier Inc. All rights reserved.

Introduction

Plants have developed a highly adaptable mechanism in order to protect themselves against exogenous RNAs such as viroids. This mechanism, called post-transcriptional gene silencing (PTGS), is induced by the presence of either double-stranded RNAs (dsRNA) or structured single-stranded RNAs (Eamens et al., 2008). These RNAs are cleaved into small interfering RNAs (siRNAs) 21 to 26 nucleotides (nt) in length by an RNase III-like enzyme, more specifically a Dicer-like enzyme (DCL). siRNAs are incorporated into RNA-induced silencing complexes (RISC) and subsequently guide the sequence specific degradation of target mRNAs. In addition, siRNAs can serve as primers for an RNA-dependent RNA polymerase (RdRp), thereby causing an amplification phenomenon. Based on the detection of siRNAs of 21 to 25 nt in length, several viroids have been shown to induce PTGS (Itaya et al., 2001, Papaefthimiou et al., 2001, Martinez de Alba et al., 2002, Landry et al., 2004, Martin et al., 2007). Moreover, it has been suggested that the production of siRNAs from viroids could explain how these pathogens cause symptoms in their hosts (Wang et al., 2004).

In the case of peach latent mosaic viroid (PLMVd), characterization using wheat germ extract shed light on the silencing mechanism. First, it was shown that the long hairpin structure implicated in the replication of the viroid has the ability to preferentially trigger DCL enzyme(s) activity(ies), a much more likely situation than the highly improbable formation of a double-stranded RNA structure including PLMVd strands of both plus (+) and minus (−) polarities (Landry and Perreault, 2005). Second, these experiments showed that the RdRp could use siRNA as primers for the production of dsRNAs that should be susceptible to DCL activity, thereby leading to amplification of the RNA silencing phenomenon as well as propagation of the PLMVd derived small RNA species (Landry and Perreault, 2005). Though these experiments were instructive, they were based on in vitro enzymatic assays, and, consequently, needed to be confirmed in vivo. With the goal of better understanding this silencing mechanism, a sequencing effort of the siRNAs associated with PLMVd infection in peach leaves was performed.

Results and discussion

Leaves were collected from one healthy and one PLMVd-infected peach trees from both the west coast of Canada and the north of Tunisia (Table 1). After RNA extraction, PLMVd infection was confirmed by both RT-PCR amplification and Northern blot hybridization using a full-length PLMVd probe of either (+) or (−) polarity as described previously (data not shown) (Bussière et al., 1999, Fekih Hassen et al., 2007). It is known that PLMVd replicates in chloroplasts via a symmetric rolling circle mechanism in which there are mainly linear monomers of both polarities, almost no detectable multimeric strands and only a small amount of circular monomers present (Bussière et al., 1999). Equivalent banding patterns were obtained for the samples derived from the two infected peach trees. The content, in

Abbreviations: DCL, Dicer-like; dsRNA, double-stranded RNA; ncRNA, non-coding RNA; PLMVd, peach latent mosaic viroid; PSTVd, potato spindle tuber viroid; PTGS, post-transcriptional gene silencing; RdRp, RNA-dependent RNA polymerase; RISC, RNA-induced silencing complex; siRNA, small interfering RNA.

⁎ Corresponding author. Fax: +1 819 564 5340.
E-mail address: Jean-Pierre.Perreault@pusherbrooke.ca (J.P. Perreault).

1 These authors have equal contributions.

0042-6822/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.11.008
molecules were reverse transcribed, PCR ampli-
to 26 nt in length were cut out of the gel and the RNA extracted. After
possessing electrophoretic mobilities that corresponded to RNAs of 18
samples were fractionated on denaturing 15% PAGE gels, the bands
more speci
characterization of the small ncRNA
Alba et al., 2002, Landry et al., 2004).

Appropriate bands were detected only in the RNA samples isolated
from infected leaves, in agreement with previous reports (Martinez de
Alba et al., 2002, Landry et al., 2004).

Table 1
Number of ncRNA by RNA sample

<table>
<thead>
<tr>
<th>Country</th>
<th>State</th>
<th>Number of total small RNAs</th>
<th>Number of unique small RNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tunisia</td>
<td>Healthy</td>
<td>189</td>
<td>165</td>
</tr>
<tr>
<td>Tunisia</td>
<td>Infected</td>
<td>156</td>
<td>156</td>
</tr>
<tr>
<td>Canada</td>
<td>Healthy</td>
<td>128</td>
<td>112</td>
</tr>
<tr>
<td>Canada</td>
<td>Infected</td>
<td>142</td>
<td>128</td>
</tr>
</tbody>
</table>

Characterization of the small ncRNA

In order to analyze the composition in terms of the small ncRNAs,
more specifically that of the PLMVd-associated siRNAs, the RNA
samples were fractionated on denaturing 15% PAGE gels (Fig. 1A).
Appropriate bands were detected only in the RNA samples isolated
from infected leaves, in agreement with previous reports (Martinez de
Alba et al., 2002, Landry et al., 2004).

![Fig. 1](image)

In order to analyze the composition in terms of the small ncRNAs,
more specifically that of the PLMVd-associated siRNAs, the RNA
samples were fractionated on denaturing 15% PAGE gels, the bands
possessing electrophoretic mobilities that corresponded to RNAs of 18
to 26 nt in length were cut out of the gel and the RNA extracted. After
the addition of linkers to the ends of the purified ncRNAs, the resulting
molecules were reverse transcribed, PCR amplified, multimerized and
finally cloned into pGEM-T vector according to the procedure
developed for microRNA sequencing (Lau et al., 2001). A total of 615
unambiguous sequences (i.e. without any unspecified nucleotides)
were obtained (see Table 1). It is not clear if the sequences found more
than once represented a higher cellular abundance, or were the result
of technical artifacts. In order to prevent any erroneous interpreta-
tions, all species were considered only once per sample in all
subsequent analyses. This yielded a total of 561 different small RNA
species, meaning that 9% of the original sequences were repeated (see
Supplementary data).

Initially, ncRNAs were analyzed as a function of their size. In the
samples derived from the healthy trees the 21 nt RNA species were
found to be less abundant than those of 24 nt (19.7% versus 33.3%,
respectively). Conversely, the 21 nt ncRNAs were slightly more
abundant than the 24 nt ones in the samples derived from infected
leaves (29.2% versus 26.8%) (Fig. 1B). This variation in the ncRNA
levels was a good indication that the antiviral RNA silencing process,
which produces 21 nt long ncRNAs, was active in PLMVd-infected
leaves.

Sequence homology searches were performed for each ncRNA
found using the available Arabidopsis thaliana plant databases since the
Prunus persica sequence is incomplete. Additional homology
searches were performed using both the available viroid subviral and
miRNA databases. In all cases, the sequences were considered to be
homologous when three or less mismatches existed between an
ncRNA and a given reported sequence (Fig. 1C). When sequence
homologies were found for more than one RNA species, the higher
homology was considered for unique assignment, if possible. Briefly,
most of the ncRNAs retrieved do not have any known homologous
sequences reported (i.e. >70%), reflecting the fact that to date none of
the peach genomes (i.e. nuclear, plastid and mitochondrial) have been
sequenced. Interestingly, we noted that the proportion of the rRNA
sequences was relatively low regardless of the origin of the samples.
This low proportion of rRNA is a good indication of the quality of the
samples. If random degradation or contamination occurred during the
preparation of the ncRNA libraries, a higher ratio of rRNA would be
expected due to the fact that rRNA represents >80% of the total cellular
RNA. We also detected a small but significant proportion that
possessed homology to the mRna and miRNA fractions, in addition
to the fact that ~20% of the ncRNA sequences possessed homology

with reported DNA sequences (data not shown). A certain proportion
of the small RNA sequences (~10%) exhibited homology to already
reported sequences, suggesting conservation of small RNA species
between organisms (Fig. 1C). Some sequences belong to two clusters,
resulting in a total of 115% fractions when all are added up. For
example, all miRNA were also part of the small RNA cluster. This type
of result is typical of several other studies (e.g. see Ref. (Martin et al.,
2007)). Importantly, no PLMVd homologous sequences were retrieved
from the samples derived from the healthy leaves supporting, once
again, the integrity of the samples and confirming the results from the
initial Northern blot hybridizations (Fig. 1A). Conversely, 20.5% of the
ncRNA sequences retrieved from the infected samples were homo-
logous to PLMVd sequences, confirming that RNA silencing was
triggered by the viroid infection (i.e. 60 out of 293 ncRNA obtained from
the infected samples). This percentage is higher than those reported for
both Potato spindle tuber viroid (PSTDV) and Citrus exocortis viroid (<5%),
two viroids that accumulate in the nucleus (Martin et al., 2007, Itaya et al., 2007). Finally, no significant bias was

![Fig. 1](image)
observed in these analyses in terms of the origin of the sample (i.e. Tunisia or Canada).

**PLMVd specific siRNAs**

The majority of the PLMVd-derived siRNA sequences (75%) were perfectly complementary to reported PLMVd variants. Only 20% of these siRNAs exhibited a single point mutation and only 5% two mismatches. Because the level of homology between the siRNAs and the PLMVd variants is very high, it is unlikely that the association between these species is a mistake. These siRNAs were primarily 21 nt in size (61.3%), corresponding to the action of the DCL1 enzyme that is involved in both the antiviral silencing as well as the miRNA production (Eamens et al., 2008). The siRNAs of 20 and 22 nt corresponded to 8.1% and 19.4%, respectively, of the siRNA population. The residual fraction (11.2%) included siRNAs of 19, 23, 24 and 26 nt. Interestingly, a significant bias for the presence of a U residue at the 5′-end (i.e. 35.5% of PLMVd like siRNA), as it is expected for siRNAs (Vaucheret, 2006), was observed. Finally, the presence of a phosphate at the 5′-end of the ncRNA, produced by DCL enzymes, was confirmed by the detection of a band shift in Northern blot hybridization if the samples were treated with alkaline phosphatase prior to electrophoresis (data not shown).

The collection of PLMVd-specific siRNAs was further analyzed in order to clarify exactly how the silencing takes place in the plant. Added together, these siRNAs contain 1368 nt, 3.9 times the length of the model PLMVd genome (338 nt). There are 35 and 25 siRNAs complementary to the PLMVd strands of (+) and (−) polarities, respectively (see Fig. 2A). It has been shown that the PLMVd strands of (+) polarity are slightly more abundant in cells than are their counterparts of (−) polarity, regardless the conformation (i.e. circular and linear monomers as well as linear multimers) (Bussière et al., 1999). An average of 3 (+) strand molecules for every 2 (−) strands has been estimated. Therefore, the slightly higher abundance of siRNA of (+) polarity observed (i.e. 58.3% compared to 41.7% for (+) and (−) polarities, respectively) likely reflects the difference in PLMVd strand abundance, indicating that strands of both polarities similarly trigger DCL enzymes. To date, there has been no DCL enzyme reported to be located within the chloroplast, consequently, this observation constitutes the first evidence that PLMVd strands of both polarities are present at some point in the cytoplasm as this is required for the RNA silencing to occur.

**PLMVd conformers as putative substrates for DCL enzymes**

*In vitro* experiments using a wheat germ extract have shown that the long P11 stem of PLMVd can trigger the DCL enzyme(s) (Landry and Perreault, 2005). However, the amplification phenomenon observed to be associated with the RNA silencing mechanism, and based on the RdRp activity, should produce a similar coverage of the PLMVd genome in terms of siRNA species. Surprisingly, the alignment of the siRNAs along the PLMVd structure showed that this was in fact not the case (Figs. 2A and B). Instead, the siRNAs appeared to be concentrated around the P2–P3 stems, the P5–P7 stems and the lower strand of the P11 stem (bold regions in Fig. 2B). While the P11 is sufficiently long, most likely the P2–P3 and P5–P7 stems have to stack together in order to form the 21 base pair helical region required for recognition by the DCL enzyme (Vaucheret, 2006). Conversely, the regions including smaller hairpins (P1, P2, P4, P8, P9 and P10) were less represented, in accordance with the dsRNA prerequisite for DCL activity. Similarly structured regions of PSTVd were shown to be preferentially hydrolyzed by the DCL enzymes (Martin et al., 2007). Differential accumulation of siRNAs for a virus has also been reported (i.e. the Cymbidium ringspot virus) (Molnar et al., 2005). Surprisingly, the upper strand of the P11 stem was less represented for both polarities, even thought it respects the requirements for DCL activity. The reason for this phenomenon remains unclear.

Further analysis of the PLMVd-associated siRNAs revealed that there are four species that overlap the hammerhead cleavage site of (+) polarity, and only one that ends precisely at the cleavage site (see Fig. 2). Localization of the observed PLMVd-associated siRNAs on the PLMVd molecule. (A) Linear representation of the PLMVd genome upon which all of the RNAs homologous to PLMVd are represented. Paired regions (P) in the secondary structure (see panel B) are shown in black, while single-stranded regions (i.e. bulges and loops) are shown in white. Due to space constraints only some paired regions are indicated. The siRNAs complementary to the PLMVd strands of (+) polarity are indicated above the genome (black lines), while those of (−) polarity are indicated below the genome (grey lines). The siRNAs covering the cleavage site are numbered 1 to 5, while those forming RNA duplexes with a 2 nt overhang at the 3′-end are labeled 6, 6′ and 7, 7′. (B) Schematic secondary structure of PLMVd according to Bussière et al., 2000. The bold lines indicate sequences covered by at least 7 siRNAs. (C) Sequences of the siRNAs covering the hammerhead cleavage site of (+) polarity. In (A), (B) and (C) arrows indicate the hammerhead cleavage sites of both polarities. (D) Sequences of pairs of siRNAs forming perfect Dicer products.
Figs. 2A and C). The cloned siRNAs were most likely protected by proteins implicated in the RNA silencing mechanism (e.g. DCL and Argonaute). Thus, their molecular stabilities should be relatively similar and, consequently, there should be no bias in terms of stability that would explain the finding of more abundant siRNA species. Three methods that explain the production of the siRNA overlapping the cleavage site can be envisioned. The first possibility is that multimeric strands of both polarities interact together, forming a long dsRNA substrate for the DCL enzyme. However, if this is indeed the method, there should be no bias in terms of the PLMVd region susceptible to DCL activity, which is not the case. The second possibility is that the dsRNA substrates for DCL activity resulted from RdRp activity using an siRNA derived from a previous PLMVd cleavage. Again, the non-uniform distribution of siRNAs is most likely an indication that this is not the explanation. The third possibility, and the most probable one, is that the PLMVd circular conformers are the preferred substrate for the DCL enzyme. The high abundance of siRNA covering the hammerhead cleavage site (4 out of 35) support this idea, especially when one considers the results from a previous quantitative analysis of the accumulation of PLMVd conformers showed that there is an average ratio of 1:11 of circular to linear strands of (+) polarity that accumulate in the leaves (Bussière et al., 1999). If the targeting of linear and circular conformer was equal, 1 siRNA coming from circular molecule would be expected for every 11 siRNAs coming from linear one. Experimentally 4 siRNA out of 35 are derived from the (+) circular genome of PLMVd, a slightly higher ratio then expected. It may be an indication that the circular conformers are selectively chosen for transport within the cytoplasm and that, once there, they become substrates for the DCL. However, it is improbable that only circular conformers are transported within the cytoplasm. The observation of an siRNA that terminates precisely at the hammerhead self-cleavage site suggests that the linear conformer is also targeted, since the cleavage of the circular one would require hydrolysis of the unusual 2′–5′-phosphodiester bond found at the self-ligation site (Côté et al., 2001) by the DCL enzyme, a highly unlikely scenario. Thus, it seems that both PLMVd circular and linear conformers of (+) polarity are present in the cytoplasm, but with a predominance of the circular conformer. This hypothesis cannot receive physical support from the analysis of the siRNAs originating from the (−) polarity PLMVd strands as no PLMVd-associated RNA was produced from the (−) strands that either overlapped with, or terminated precisely at, the hammerhead cleavage site. This might result from the fact that the number of siRNAs from the upper strand of PLMVd was limited for unknown reasons as mentioned above. Clearly, additional studies will be required in order to clarify this unexpected accumulation of fragments covering the hammerhead cleavage site of (+) polarity.

Finally, we wondered if the collection of sequences contains any pair(s) of either partially or perfectly complementary RNAs that could form dsRNA harboring 2 nt overlaps on their 3′ end corresponding to an ideal DCL cleavage. Two such pairs were found (Fig. 2D), one in the P11 stem and the other in the P2–P3 stems. Both cases were perfectly homologous, indicating that the DCL substrate could result from the RdRp activity. Further biochemical experiments should help in solving this inconsistency. In opposition to what was suggested previously, this observation supports the involvement of the RdRp activity in the RNA silencing amplification. Of course, the RNAs forming the perfect duplexes could originate from two different cleavage activities and been found solely by chance. However, it is surprising that no duplex coming from the P11 stem and possessing the corresponding mismatches, was not also found.

Concluding remarks

This sequencing effort and the analysis of the PLMVd-associated siRNAs reveal several aspects of the silencing defense mechanism in plants. Specifically, PLMVd strands of both polarities are substrates for the DCL enzymes that preferentially recognize the most structured region of the viroid. Moreover, considering that DCL1 is associated with antiviral RNA silencing, and that it is found in the cytoplasm, the data presented here indicate that although PLMVd strands of both polarities are transported within the cytoplasm, the circular conformers are significantly more efficiently transported. It is not known if it is the presence of a sequence that could be juxtaposed with the circular conformer ligation site, or the presence of the unusual 2′,5′-phosphodiester bond at that position, that provides the selective advantage to the circular conformers. Furthermore, it is not impossible that the circular conformers adopt a different structure, as compared to the linear ones, which would favor their exportation. This collection of ncRNAs sequences sheds light on several aspects of silencing of a viroid; however, it also raises several questions. The availability of high-throughput sequencing technology will certainly be an important tool in answering several of these intriguing questions.

Materials and methods

RNA extraction and Northern blot hybridization

Total RNA (40 µg) was extracted from both infected and healthy peach leaves grown on the west coast of Canada (Sidney, British Columbia) and in the north of Tunisia using TRIzol as recommended by the manufacturer (Invitrogen). These RNA samples were quantified by UV spectroscopy, and their quality verified by 1% agarose gel electrophoresis. The replication of PLMVd was verified by both RT-PCR amplification and Northern blot hybridization as described previously (Bussière et al., 1999, Fekih Hassen et al., 2007). PLMVd-associated siRNAs were detected by Northern blot hybridization as reported previously (Landry et al., 2004).

Cloning and analysis of small ncRNAs

Cloning and sequencing of the small ncRNAs was performed as described previously with only minor modifications (Lau et al., 2001, Motard et al., 2008). Briefly, RNA samples were fractionated on denaturing 15% PAGE gels. The gel bands corresponding to the region located between the 18 and 26 nt long RNA species markers were visualized by UV shadowing, excised, the RNA eluted overnight at 4 °C, ethanol precipitated and dried. Linkers were then ligated to both ends of the small ncRNA as described previously (Landry and Perreault, 2005, Lau et al., 2001). The 3′-linker (5′-PO4-5′CUGUAGATCATCA-AH2′-3′, the bold letters represent ribonucleotides, while the other positions are deoxyribonucleotides) includes a 5′ phosphate (PO4) and a 3′ NH2 in order to avoid multimerization between the monomers, while the 5′-linker (5′ATCCUGAGAUCCUGAAA-3′) did not posses any special modification. The resulting products were reverse transcribed, PCR amplified, multimerized, ligated into pGEM-T vector (Promega), transformed into E. coli XLI-Blue bacteria and the resulting recombinant plasmids purified and then inserted sequenced at the McGill Genome Center. Sequence homology searches were performed for each ncRNA found using the available A. thaliana (http://www.ncbi.nlm.nih.gov/mapview/static/MVPlantBlast.shtml?), viroid subviral (http://subviral.med.ualberta.ca/cgi-bin/home.cgi; (Rocheleau and Pelchat, 2006)), viral (http://gibgen.nig.ac.jp; (Hirahata et al., 2007) and miRNA (http://microrna.sanger.ac.uk/sequences; (Griffiths-Jones et al., 2004)) databases.

Acknowledgments

The authors would like to thank Olivier Parisi for the greenhouse work. This work was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC, grant 155219-07) of Canada. The RNA group is supported by grants from both the Université de Sherbrooke and the Canadian Institute for Health Research.
Research. J.P.P. holds the Canada Research Chair in Genomics and Catalytic RNA and is member of the Centre de Recherche Clinique Etienne-Lebel.

Appendix A. Supplementary data


References


