Characterization of plant miRNAs and small RNAs derived from potato spindle tuber viroid in infected tomato

Natalie Diermann\textsuperscript{a}, Jaroslav Matoušek\textsuperscript{b}, Jan-Hendrik Teune\textsuperscript{a}, Detlev Riesner\textsuperscript{a}, Gerhard Steger\textsuperscript{a,\textsuperscript{*}}

\textsuperscript{a}Heinrich-Heine-Universität Düsseldorf, Institut für Physikalische Biologie, Geb. 26.12.U1, Universitätsstr. 1, D-40225 Düsseldorf, Germany
\textsuperscript{b}Biological Centre AS CR v.v.i., Institute of Plant Molecular Biology, Braníčkova 31, 37005 České Budějovice, Czech Republic

Abstract

To defend against invading pathogens, plants possess RNA silencing mechanisms involving small RNAs like miRNAs and siRNAs. Several groups have reported that also viroids—plant infectious, non-coding, unencapsidated RNA—cause the production of viroid-specific small RNAs (vsRNA), but viroids do escape the cytoplasmic silencing mechanism. On the other side viroids with only minor sequence variations can produce quite different symptoms in infected plants. This finding suggests an involvement of vsRNAs in symptom production. Here we analyze by deep sequencing the spectrum of vsRNAs induced by the potato spindle tuber viroid (PSTVd) strain AS1, which causes very strong symptoms like dwarfing, leaf distortion and necrosis upon infection of tomato plants cv Rutgers. Indeed vsRNAs mapping to the pathogenicity-modulating domain of PSTVd were present at highest rate, supporting an involvement of vsRNAs in symptom production. Furthermore, in PSTVd AS1-infected plants the accumulation of some endogenous miRNAs, which are known to be involved in leaf development via regulation of transcription factors, is suppressed. The latter finding supports the hypothesis that a miRNA-dependent (mis-)regulation of transcription factors upon PSTVd infection causes the viroid symptoms.

\textit{Key words:} microRNAs, transcription factor, viroid pathogenicity

\textbf{Introduction}

Viroids are plant pathogens that consist of a single-stranded circular RNA of 246–475 nucleotides (nt) in length. Unlike viruses viroids do not code for any protein and thus do not possess a capsid. Nevertheless viroids replicate autonomously in the host cell (for reviews see Hadidi et al., 2003; Tsaigris et al., 2008; Ding, 2009). To utilize the host’s biochemical machinery viroids have to present appropriate signals that are coded for by their sequence and structure. In the following we concentrate on viroids in their sequence and structure. In infected tomato plants cv Rutgers viroid genomic RNAs (Itaya et al., 2001; Ma-

Whereas the molecular mechanism of viroid replication are fairly well known, only vague models could be proposed for pathogenesis that accounted for the correlation between sequence variations and symptoms. Based on the first five PSTVd strains Schnöller et al. (1985) suggested that decreasing thermodynamic stability of the P domain leads to increasing virulence of the corresponding isolate. With increasing numbers of strains this model was rendered untenable (Owens et al., 1996). A further model correlated the three-dimensional bending of the P domain with symptoms (Owens et al., 1996; Schmitz and Riesner, 1998). This model, however, does not extend to the most severe PSTVd strain, AS1 (Matoušek et al., 2007).

The detection of viroid-specific small RNA (vsRNA) species of 21–24 nt in length that accumulate during viroid propagation (Papaefthimiou et al., 2001; Itaya et al., 2001; Markarian et al., 2004; Wang et al., 2004; Machida et al., 2007) might point to an involvement of vsRNA in post-transcriptional gene-silencing (PTGS) processes (for reviews see Flores et al., 2005; Gómez et al., 2009). Normally it is assumed that virus-derived small interfering RNAs counteract virus replication by silencing (for review see Shukla et al., 2008). With viroids, however, a higher accumulation of vsRNA was found to be associated with higher accumulation of viroid genomic RNAs (Itaya et al., 2001; Ma-

Preprint submitted to Virology August 20, 2009

\textsuperscript{*}Corresponding author. FAX: +49 211 81 15167 Phone: +49 211 81 14597
\textit{Email addresses:} diermann@biophys.uni-duesseldorf.de (Natalie Diermann), nat.j@centrum.cz (Jaroslav Matoušek), teune@biophys.uni-duesseldorf.de (Jan-Hendrik Teune), riesner@biophys.uni-duesseldorf.de (Detlev Riesner), steger@biophys.uni-duesseldorf.de (Gerhard Steger)
Figure 1. Secondary structure of PSTVd strains and their symptoms in tomato plants cv Rutgers. A: The secondary structure of PSTVd consists of a linear, rod-like arrangement of short helices and loops. The size of dots connecting base pairs is based on thermodynamic prediction of a consensus structure for 45 PSTVd strains; the sequence is given for the PSTVd variant Intermediate (Gross et al., 1978). The five homology domains of pospiviroids are marked as proposed by Keese and Symons (1985): terminal left and right (TL, TR), pathogenicity-modulating (P), central conserved (C), and variable (V) domain. B: Different strains of PSTVd, which differ only by few mutations in the P domain (see nucleotides on grey background in the insets), produce different symptoms in infected plants. QFA (Gruner et al., 1995) is a mild strain, whereas AS1 (Matoušek et al., 2007) is a very severe (“lethal”) strain producing symptoms up to necrosis of stem and leaves. The left-most plant is mock-inoculated; all plants are of identical age 16 days post inoculation (d. p. i.).
synthetic (+)-stranded PSTVd by using a recombinant human
dicer enzyme. Most vsRNAs mapped to lower TR and upper
TL and V domains of the (+)-strand PSTVd RNA; only in
vitro vsRNAs were generated from the P domain. These re-
sults would point against involvement of vsRNA in symptom
generation in PSTVd-infected plants.

Machida et al. (2007) sequenced 89 vsRNAs from a single
PSTVd-infected tomato leaf. In contrast to the finding of Itaya
et al. (2007) many vsRNAs were derived from the (−)-strand
(mostly lower TR, lower C, and upper TL to P domains); the
(+)-vsRNAs mapped mainly to upper TL to P, upper V, lower
TR and lower C domains. Only two of the (+)-stranded and
eight of the (−)-stranded vsRNAs overlapped with the upper or
lower P domain; that is, only a minority of vsRNAs could be
related to symptom generation.

In a similar study on tomato plants infected with citrus exo-
cortis viroid (CEVd, family Pospiviroidae) Martín et al. (2007)
found again that most vsRNAs were derived from (+)-stranded
CEVd and there from upper TL, upper TR, and lower V and
C domain. In addition they tested the hypothesis that viroid
infection might alter miRNA biogenesis and/or function, but
found no clues pointing in this direction.

The present study is based on our earlier finding with the
“lethal” PSTVd strain AS1 (Matoušek et al., 2007) that ele-
vated levels of vsRNA are associated with production of strong
symptoms. Here we want to study especially the interdepen-
dence of vsRNA production and symptom generation. For this
we analyze in particular in tomato plants cv Rutgers: the quan-
tities and origins of vsRNAs by hybridization with different
PSTVd RNA fragments of both polarities; the origin of par-
tially purified vsRNAs by Solexa sequencing; the difference in
small RNAs present in mock-infected plants in comparison to
PSTVd-infected ones; and, given differences in the presence of
miRNAs between infected and non-infected plants, potential in-
fluences on host metabolism. Especially in the first three tasks
we took great care not to lose any vsRNAs during the necessary
purification steps; most vsRNAs critical for symptom genera-
tion might be hybridized to their (unknown) target molecules.
Indeed we found that the majority of vsRNA map to the P do-
main and that the level of several endogenous miRNAs is al-
tered upon infection with PSTVd AS1.

Results

Northern blot analysis

For enrichment of small RNAs from a total RNA extract
of tomato plants, we tried first to follow standard protocols
(Schmitz and Riesner, 2006; Itaya et al., 2007) using differen-
tial precipitation with 5 % polyethylene glycol (PEG). RNAs from
supernatant and pellet were separated by denaturing PAGE,
blotted, and hybridized with 5’-[32P]-labeled full-length (+) or
(−)-PSTVd RNA; autoradiographs are shown in Fig. 2. Obvi-
ously the pellet fraction, which should only contain long RNAs,
contains a small but significant amount of small RNAs, espe-
cially in case of the RNA extracted from AS1-infected plants.
Such a loss might be critical because pathogenesis-related small
RNAs could particularly be hybridized to their target RNAs.

Thus in the following experiments we resorted to denaturing
gel electrophoresis and gel elution for the enrichment of small
RNAs.

For analysis of a potential time course in levels of vsRNAs
and to get some insight into the mapping of vsRNAs to the
PSTVd sequence we used Northern blot analysis with eight dif-
f erent hybridization probes: these probes correspond to quarts
of full-length PSTVd sequence in (+) and (−)-orientation (for
sequences see Table 1 left). As controls we used small syn-
thetic RNAs that map close to the center of the quarter probes
(for sequences see Table 1 right). RNA extracts were made
from tomato plants (S. lycopersicum cv Rutgers) infected with
two different PSTVd strains: QFA is a mild strain and AS1 a
very severe one (see Fig. 1).

RNAs extracted from mock-inoculated plants do not show
any signal in the Northern blot analysis (see Fig. 3). In con-
trast the viroid-infected plants show clear signals with all (+)–
as well as (−)-probes. All vsRNAs are seen as a single band
of about 21 nt at all time points. This differs from the observa-
tion by Machida et al. (2007) who observed two different size
classes of vsRNA (21 and 24 nt). We could not observe a uni-
form time course: the QFA-vsRNAs seem to be at higher levels
at days 30 and 37 than at days 23 or 44; but the opposite is
visible for AS1-vsRNAs. The radioactive signals of vsRNAs
hybridized to the different quarter probes are quite comparable
to each other, but we needed different exposition times between
1 and 5 days for the different Northern blots; blots at 1 day ex-
position time are shown in supplemental Fig. S1. On basis of
the signals obtained with the small RNA controls, these differ-
ences in exposition times seem, however, to be related more to
the hybridization efficiency of the different quarter probes than
to different levels of vsRNAs mapping to the different quarts
of the PSTVd sequence in both polarities.

These results do not allow for firm conclusions on a (equal
or unequal) distribution of vsRNAs to different regions of the
PSTVd genome and their polarity. To allow for more precise
insight into the origin of the vsRNAs we decided to do a high-
throughput sequencing of small RNAs.

Figure 2. Co-precipitation of small and long RNAs in a PEG pre-
cipitation. An enrichment of small RNAs from total RNA extracts
of PSTVd QFA- and AS1-infected plants was tried by precipitation
using 5 % PEG and 0.5 M NaCl. The four central lanes contain syn-
thetic RNAs of 20 to 24 nt length as size markers. The left blot was
hybridized with (+)-stranded full-length PSTVd RNA, the right with
(−)-stranded RNA.
Figure 3. Northern blot analysis of vsRNAs in a total RNA extract from PSTVd QFA- and AS1-infected tomato plants. RNA was hybridized with radioactively labelled PSTVd-quarter probes marked on the left (for sequences see Table 1 left); the area between 20 and 26 nt is shown. Five ng of full-length PSTVd RNA (FL) of both orientations was applied to the first two lanes, which show no signal in this size range. In the next 3×4 lanes 10 µg total RNA from mock-treated, QFA- and AS1-infected plants from four different time points (23, 30, 37, and 44 d.p.i.) are applied. In the right eight lanes 15 ng small synthetic RNAs (for sequences see Table 1 right) were used both as length standards and positive or negative controls, respectively. Exposure times were, from top to bottom, 2, 3, 5, 2, 5, 1, 2, and 5 days, respectively.

Sequencing of small RNAs

Gel-purified small RNAs from mock-inoculated and PSTVd AS1-infected plants were subjected to deep sequencing and bioinformatic analysis as described in the Materials and Methods section.

vsRNAs

As is obvious from Fig. 4, vsRNAs with (+)-PSTVd sequence dominated by far over those with (−)-PSTVd sequence (4007 versus 447 sequences). (+)- as well as (−)-vsRNAs had a mean length of 20.4 nt. Most (−)-vsRNAs mapped to the border between upper C and V domain. The majority of (+)-vsRNAs mapped to five different regions: upper TL domain (f. e. 192× PSTVd sequence 6–26, 73× 18–37), lower TR (f. e. 57× 189–210), between lower V and C (f. e. 33× 228–248), lower C (f. e. 65× 259–279), lower P (f. e. 140× 294–313), and between lower P and TL domain (f. e. 288× 302–322, 151× 302–321). That is, vsRNAs mapping to the lower P and between lower P and TL domain were represented at highest rate.

Non-PSTVd-related small RNAs

The data sets on small RNAs from mock-inoculated and PSTVd-infected tomato plants contained a large amount of fragments of ribosomal RNAs and tRNAs, which might simply be degradation products. For example, 3557 (0.29 %) and 1335 (0.40 %) fragments, respectively, were present mapping to tomato 17S rRNA (GenBank AC X51576).

A high number of fragments mapped to chromosome 2 of tomato as well as to chloroplastic NADH dehydrogenase ND2 subunit [35023 (2.86 %) and 12537 (3.78 %) fragments, respectively]; these double hits are due to the transfer of plastid-derived DNA to plant nuclear genomes (Huang et al., 2005).

Furthermore, many mature miRNAs were present in both

Figure 4. Position of small RNAs from the PSTVd-AS1 genome found by Solexa sequencing. The x-axis represents the PSTVd-AS1 genome. The curve above the x-axis (positive numbers) represents vsRNAs with (+)-sequence, below (negative numbers) with (−)-sequence. The most prominent vsRNAs (≥20× identical sequence) are marked by black bars. Thickness of bars is proportional to amount of vsRNAs with this sequence; for example, the most common vsRNA (with PSTVd sequence 302–322) occurs 288 times. The five homology domains of pospiviroids are marked and labelled in grey (see Fig. 1A).

Non-PSTVd-related small RNAs

The data sets on small RNAs from mock-inoculated and PSTVd-infected tomato plants contained a large amount of fragments of ribosomal RNAs and tRNAs, which might simply be degradation products. For example, 3557 (0.29 %) and 1335 (0.40 %) fragments, respectively, were present mapping to tomato 17S rRNA (GenBank AC X51576).

A high number of fragments mapped to chromosome 2 of tomato as well as to chloroplastic NADH dehydrogenase ND2 subunit [35023 (2.86 %) and 12537 (3.78 %) fragments, respectively]; these double hits are due to the transfer of plastid-derived DNA to plant nuclear genomes (Huang et al., 2005).

Furthermore, many mature miRNAs were present in both

Figure 5. Number of mature miRNA sequences present in Solexa data from AS1-infected and mock-inoculated plants, respectively. The number of sequences from AS1-infected plants was scaled by 3.69 to compensate for the lower number of total sequences. The factors next to the bars give the ratio of miRNAs from AS1-infected to mock-inoculated samples.
PSTVd RNA as a miRNA precursor?

One possibility for generation of the vsRNAs sequenced in this and other projects (Itaya et al., 2007; Machida et al., 2007) is a miRNA-like pathway; another would be a ta-siRNA-like biogenesis (Gómez et al., 2009). To get some insight into the capability of PSTVd to mimic a miRNA precursor and to identify potential miRNA locations in PSTVd RNA we developed a program YAMP (Teune, Steger, unpublished; for some details see Materials and Methods) that uses a statistical model to discriminate a miRNA precursor from all other RNAs and does not rely on prior knowledge of a miRNA target or comparative genomics. Note that the program is highly sensitive and specific in detection of a miRNA/miRNA* position but is unable to decide upon which is the miRNA in the miRNA/miRNA* duplex.

For an RNA with PSTVd sequence from positions 263–359/1–96, which is one of the structural elements present during processing of (+)-strand replication intermediates to circles (Baumstark et al., 1997), YAMP predicted a further miRNA/miRNA* complex in the P domain of PSTVd (positions 288–307, 293–311, 294–312, 295–313, 297–315, and 300–318; see Fig. 6, lines marked “miRNA prediction”); for an RNA from positions 103–255, which is also a structural element during processing, YAMP predicted a further miRNA/miRNA* complex in the TR domain (149–168, 150–169, 151–170, and 152–171). For a nearly full-length, linear PSTVd (2–358) YAMP predicted identical positions for complexes. For half PSTVd molecules (either 1–180 or 181–369) YAMP predicted no miRNA; this is as expected because these half-molecules are not able to fold back into a pre-miRNA-like structure.

Possible targets of vsRNAs

According to the hypothesis that vsRNAs are involved in symptom generation in PSTVd-infected plants, at least one of these vsRNAs should be complementary to a host mRNA as expected because these half-molecules are not able to fold back into a pre-miRNA-like structure. Possible targets of vsRNAs E

For an RNA with PSTVd sequence from positions 263–359/1–96, which is one of the structural elements present during processing of (+)-strand replication intermediates to circles (Baumstark et al., 1997), YAMP predicted a further miRNA/miRNA* complex in the P domain of PSTVd (positions 288–307, 293–311, 294–312, 295–313, 297–315, and 300–318; see Fig. 6, lines marked “miRNA prediction”); for an RNA from positions 103–255, which is also a structural element during processing, YAMP predicted a further miRNA/miRNA* complex in the TR domain (149–168, 150–169, 151–170, and 152–171). For a nearly full-length, linear PSTVd (2–358) YAMP predicted identical positions for complexes. For half PSTVd molecules (either 1–180 or 181–369) YAMP predicted no miRNA; this is as expected because these half-molecules are not able to fold back into a pre-miRNA-like structure.

Possible targets of vsRNAs

According to the hypothesis that vsRNAs are involved in symptom generation in PSTVd-infected plants, at least one of these vsRNAs should be complementary to a host mRNA and lead to slicing or repression of this target. A simple BLAST search for sequences of Solanaceae, to which the typical PSTVd hosts belong, with partial homology or complemen-
tarity to PSTVd yielded no longer fragments than expected by chance. Next, we wanted to analyze a complete plant genome for complementarity to PSTVd. Because the tomato genome, as the standard laboratory host genome, is only partially available (and not annotated), we selected the Arabidopsis thaliana genome; this plant is not a host, but at least supports replication of viroids (Daròs and Flores, 2004; Matoušek et al., 2004). The search was performed using PatScan allowing for hits with no more than one mismatch, one insertion and one deletion. Depending on the PSTVd strain used, about 100 to 200 hits were obtained. About 40% of hits were located in regions annotated as protein-coding in the Arabidopsis genome. In the PSTVd sequences up to 90% of hits are located in the upper TL to upper P domains and in the lower P domain; these regions are marked by arrows in Fig. 6 (see lines marked “PatScan hits”). A statistical significance of this result is not easy to evaluate due to the nucleotide bias of the P domain, which contains an oligopurine stretch in the upper and an oligo-pyrimidine stretch in the lower part.

The obvious down-regulation of miR159, miR396, miR319, and miR403 in PSTVd-infected plants might suggest a direct interaction of vsRNAs with the corresponding pri-miRNAs. By NCBI-BLAST searches we identified possible genomic regions for miR319 (GenBank ACs DU051444, AC217334, AC217248; compare Zhang et al., 2008) and miR403 (AC234540, BI176266, CZ961341; compare Sunkar and Jagadeeswaran, 2008). In these sequences we searched—by NCBI-BLAST, PatScan, or RNAHybrid (Rehmsmeier et al., 2004)—for significant regions of complementarity to PSTVd. The region of complementarity between pri-miR403 and PSTVd with lowest free energy is shown in Fig. 7; this local complementarity has favorable energy but neither looks like a typical example of a miRNA/target interaction nor is this PSTVd fragment among the prominent vsRNAs (see Fig. 4). A similar region of complementarity was not found for pri-miR319.

### Discussion

Some findings, from this work and from the literature, concerning PSTVd and its relation to PSTVd-derived small RNAs are summarized schematically in Fig. 6:

First, all experimentally derived information shows that vsRNAs are not equally distributed over the PSTVd genome but are derived only from particular regions. This finding supports a miRNA-like generation of vsRNA. The same hypothesis is supported by the YAMP predictions, which accepts PSTVd half and full-length structures as pre-miRNAs. In case of a ta-siRNA-like generation of vsRNA—as proposed by Gómez et al. (2009) based on the dependence of hop stunt viroid-induced symptoms in Nicotiana benthamiana on the expression of RDR6 (Gómez et al., 2008)—one would expect rather an occurrence of phased small RNAs (Allen et al., 2005).

Second, the regions of CEVd-specific small RNAs seem not to overlap fully with those of PSTVd-specific small RNAs, despite the close sequence and structural similarities between the two pospiviroids CEVd and PSTVd. It is an open question as to whether these differences are based only on different types of analysis or if they are biologically significant.

Third, some vsRNA sequences found in the three studies on PSTVd are in agreement, others show clear differences: the vsRNAs mapping to the upper TL and lower TR domains coincide to a large extent; Itaya et al. (2007) and Machida et al. (2007) found vsRNAs from the upper V and lower TL domain; Machida et al. (2007) and our study identified vsRNAs from the lower C domain. Only our study, however, identified vsRNAs from the lower P domain. This is a remarkable finding, it might be due to some subtle differences and biases in the experimental approaches. For example, Itaya et al. (2007) used a PEG precipitation for enrichment of small RNAs, which led to some losses of small RNAs at least in our hands. Furthermore, we used the PSTVd strain AS1, which induces more severe symptoms than the strains used in the other studies. Clearly this asks for a repetition of the deep-sequencing approach with at least another PSTVd strain of pathogenicity different from strain AS1.

Fourth, according to the genomic search in Arabidopsis, vsRNAs derived from upper and lower P domain would be complementary to several mRNAs; there is, however, not a clear indication to mRNAs or non-coding RNAs relevant for PSTVd infection or pathogenesis. A similar genomic search in S. lyrata to the TCP4 mRNA of A. thaliana and the TCP Lanceolate gene of tomato in a conserved motif close to their 3’ ends (see Fig. 8). Because viroid infection and pathogenesis obviously mediate down-regulation of miR319 we asked about a possible up-regulation of the TCP Lanceolate gene in tomato. Semi- quantitative RT PCR was performed using total RNA isolated from tomato cv Rutgers leaf blades, which were either mock inoculated or biologically infected with PSTVd strain AS1 showing strong symptoms (see Fig. 9). Our results strongly suggest an up-regulation of the TCP gene in leaf blades of PSTVd-infected tomatoes.
Figure 8. Nucleotide and protein alignment of a sequence section of TCP factors including the position of miR319 binding. A) Nucleotide alignment of LANCEOLATE mRNAs from *S. lycopersicum* (GenBank AC AK247219: cv MicroTom, nt 1436–1528; AC EF091571: cv M82; nt 1425–1517) and from *A. thaliana* (AC NM_112365: nt 1241–1336). Above and below the alignment are given the miR319 sequences from *S. lycopersicum* and *A. thaliana*, respectively, including complementarity to their targets. B) Protein alignment of the identical region as in A; the miR319 binding site is boxed.

*lycopersicum* cannot be carried out because its genome sequence and annotations are still missing.

Most remarkable, however, is the non-uniform ratio of specific miRNAs in the data sets from healthy plants and from the plants strongly affected by PSTVd disease (see Fig. 5). This is in contrast to the study of Martín et al. (2007) who showed that CEVd is unlikely to affect levels of miR319 and others; note, however, that with CEVd also no vsRNAs from the lower TR, TL, and P domain were identified.

The finding of reduced levels of miRNAs is consistent with the assumption that certain plant regulatory factors are involved in disease development via miRNAs (e.g. Zhang et al., 2005). The involvement of regulatory factors in viroid pathogenesis can be assumed also from other observations summarized elsewhere (for reviews see Hadidi et al., 2003; Tabler and Tsagris, 2004): symptom expression is strongly dependent on the stages of plant development; some developmental changes are tissuespecific; pathogenesis is very complex on the transcriptome level, and regulatory networks including phytohormones like ethylene and gibberellic acids are involved in pathogenesis.

The miR403 targets AGO2, which belongs to the argonaute protein family involved in all known small RNA-directed regulatory pathways (Vaucheret, 2008); however, the biological significance of AGO2 in small RNA metabolism remains unknown (Ramachandran and Chen, 2008). The miR396 belongs to stress-inducible (high-salinity-, drought-, and cold-regulated) miRNAs in *Arabidopsis* that target transcription factors involved in further regulation of gene expression and signal transduction probably functional in stress responses (Liu et al., 2008). In *Arabidopsis* miR159a and b regulate the expression of a family of seven transcription factors that includes MYB33 and MYB65; a double mutant of miR159a and b shows, besides other morphological defects, altered growth habit and curled leaves (Allen et al., 2007). Expression of miR319 in *S. lycopersicum* down-regulates lanceolate-like genes resulting in larger leaflets and continuous growth of leaf margins (Ori et al., 2007).

Up-regulation of tomato TCP could be involved in the development of pathogenesis-related symptoms, especially in leaf blade malformations and morphogenic changes. Moreover, it is rather the rule than the exception that individual transcription factors (TFs) activate or inactivate set(s) of genes by a combinatorial manner (Singh, 1998). In our previous work (Matoušek et al., 2007) we described strong stem and petiole-specific activation of a set of apoptotic nucleases involved in degeneration of the tomato vascular system in strongly stunted plants. Although some vascular system-specific tomato TFs like VSF-1...
(Torres-Schumann et al., 1996) could be interconnected to development of symptoms in relation to host miRNA “disbalance” detected in the present work, such possibility remains to be verified.

In the present study we showed that viroid infection leads to processing of viroids into specific vsRNAs and, in addition, to suppression of plant-derived miRNAs possibly leading to leaf malformation and stunting. Because only in one case the link between a vsRNA and a plant miRNA could be suggested, it will be the main task of future research—when the host genome is available—to identify the major host targets of vsRNAs.

### Materials and Methods

#### Plant material

Tomato plants (*S. lycopersicum*) cv Rutgers were used for extracting total RNA for Northern blotting as well as for high-throughput sequencing. All plants were grown in a greenhouse at a temperature of 26 to 30°C with natural light. After development of the primary leaf, plants were inoculated by using Carborundum as abrasive and 10 µl sodium phosphate buffer (pH 7.0) containing 0.5 ng circular PSTVd RNA of the respective viroid strain (crude RNA extract prepared as described in Schmitz and Riesner, 2006); QFA, Intermediate, and AS1 were used. Samples were taken after 23, 30, 37 and 44 d.p.i. for Northern blot analysis and after 30 days for sequencing. Total RNA was isolated by phenol/chloroform extraction and ethanol precipitation.

#### Northern blot analysis

For Northern blotting 10 µg total RNA per lane of each sample was applied on a 14 % (30:1) polyacrylamide gel. Five ng full-length linear PSTVd as well as 15 ng synthetic small RNAs were used as markers.

After separation the RNA was blotted by a semidry electrotblotter (Panther, Thermofisher, Dreieich, Germany) onto a positively charged nylon membrane (Amersham Hybond-N+, GE Healthcare) for 30 min at 14 V and 400 mA. Nucleic acids were covalently fixed to the membrane by irradiation with 254 nm UV light (120 mJ). Hybridization was conducted with 5× Denhardt’s solution, 5× SSC, 0.1 % SDS and 250 µg/ml hering sperm DNA at 65°C in tubes in a hybridization oven over night. Afterwards the membrane was washed 2× for 30 min with 2× SSC, 0.1 % SDS at 55°C and exposed to film at −70°C.

#### Hybridization probes and oligonucleotides

RNAs with sequence of PSTVd but about a quarter length of PSTVd RNA were used as hybridization probes. For a graphical overview of probe positions see supplemental Fig. S2; start and end positions of probes are given in Table 1 (left). The probes were transcribed from templates made via PCR (including a T7 promoter) and end-labelled with 5’-[32P] Pcp (Bruce and Uhlenbeck, 1978). For primer sequences see supplemental Table S1.

Small synthetic oligonucleotides (CureVac, Tübingen, Germany) of distinct lengths and PSTVd sequence were used as positive and negative hybridization controls as well as size markers; start and end positions of oligonucleotides are given in Table 1 (right).

### Sequencing of small RNAs

For high-throughput sequencing close to 5 µg small RNAs were necessary. For this, about 15 tomato plants were inoculated with PSTVd strain AS1; at 30 d.p.i. leaves were pooled and total RNA was separated by phenol/chloroform extraction and ethanol precipitation. We tested several recent methods for complete separation of small from longer RNAs: binding of siRNAs to RNeasy columns (Qiagen) using high isopropanol concentrations (for details see http://www.laborjournal.de/rubric/tricks/tricks/tricks/trick80.lasso), miRNeasy kit (Qiagen), separation of siRNA from longer total RNA by centrifugation plus filtration according to protocols of the manufacturers (Microcon centrifugal filter units, Millipore; Vivaspin-500, Vivaspin-500, Sartorius). The classical denaturing gel electrophoresis plus gel-elution turned out to be most efficient.

After denaturation at 95°C the RNA extracts (~6 mg) were separated on a 40 cm long, 2 mm thick 20 % (19:1) polyacrylamide gel (slot width 24 cm) with 8 M urea and 0.5× TBE at about 35 W and 50°C over night. Afterwards the relevant size region of the gel was excised and the RNA electro-eluted (HSB-Elutor, Biometra) from the low-salt gel-electrophoresis buffer into a salt sink (3 M sodium acetate, pH 6.0). To minimize loss of RNA, LoBind Tubes (Eppendorf, Germany) were used.

Ligation of 3’ and 5’ adapters and cDNA synthesis was performed by vertis Biotechnologie (Freising, Germany). Briefly, the small RNAs were first treated with tobacco acid pyrophosphatase, then poly(A)-tailed using poly(A) polymerase followed by ligation of an RNA adapter to the 5’-phosphate; first-strand cDNA synthesis was then performed using an oligo(dT)-adapter primer and MMLV-RNase H+ reverse transcriptase; resulting cDNAs were PCR-amplified to about 20 ng/µl in 17 cycles using a high fidelity DNA polymerase; the cDNAs were purified using the NucleoSpin Extract II kit (Macherey & Nagel, Düren, Germany). Solexa sequencing of the cDNA was done by GATC Biotech (Konstanz, Germany).

Processing of Solexa sequence data

By Solexa sequencing of the small cDNA from mock-inoculated and PSTVd AS1-infected tomato plants about 1.2 and 0.33 million reads, respectively, were obtained, after removal of adapter sequences from the individual sequences and deleting sequences shorter than 17 nt and containing more than 90% adenines, respectively.

The two data sets were searched for PSTVd AS1-related sequences and mature miRNA sequences, respectively, using tools from the VMATCH package (Abouelhoda et al., 2004, www.vmatch.de) and a few Perl scripts to parse their outputs. Perfect matches are summarized in paragraph "vsRNAs" and "Non-PSTVd-related small RNAs" (Fig. 4), respectively. In addition to the perfect matches, hits with one mismatch (∼0.65), with a mismatch at the 5’ end (∼0.05) or the 3’ end (∼0.15) were found (data not shown). These amounts of non-perfect matches coincide with values known from the literature; compare for example to Dohm et al. (2008). Taking these additional hits into account altered neither preferred locations of vsRNAs in the PSTVd genome nor the relative amounts of miRNAs.

The sequences used in searching were assembled from GenBank (chloroplast genomes and ribosomal RNAs from solanaceous plants), from the tRNA data base (Sprinzl and Vassilenko, 2005), and from miRBase (version 13.0; Griffiths-Jones et al., 2007).

We compared both data sets for sequences that were under- or over-represented. For this purpose we used CD-HIT-EST-2D (Li and Godzik, 2006) with options -c 95 -n 8 -r 1 -p 1 -d 40 to ensure that only sequences with more than 95% sequence identity were clustered. Sequences represented more than 100 times were inspected manually; for example sequences differing by a few terminal nucleotides were combined. Annotations of these clustered sequences were done via NCBI-BLAST in the “nucleotide collection” restricted to Solanaceae.

Search for PSTVd-related sequences in the genome of A. thaliana

To search for potential targets of vsRNAs, the sequence of PSTVd strain AS1 was cleaved into 359 overlapping fragments of length 23 nt; these were searched for with PATSCAN (Dsoouza et al., 1997) in the Arabidopsis genome. Complementary fragments with maximally one mismatch, one insertion and/or one deletion were counted as matches.

Prediction of miRNAs in PSTVd RNA by YAMP

To identify potential miRNA locations in PSTVd RNA we were in need of a program that relies neither on comparative genomics nor on prior knowledge of a miRNA target (as other programs do; f. e. Adai et al., 2005; Dezulian et al., 2006). The program YAMP (Teune, Steger, unpublished) uses a series of filtering methods followed by a statistical model to discriminate a miRNA precursor from all other RNAs.

The first step before classifying potential miRNA candidates involves the calculation of secondary structures and their conversion into an alignment-like format: the introduction of a gap-symbol caused by asymmetrical loops and bulge loops results in either base pairs, non-pairing nucleotides in loops, and nucleotide/gap pairs in asymmetrical loops. This representation of the secondary structure is used during the filter steps and classification procedure.

The thresholds for the filter steps and statistical model are based on a set of true positive sequences (151 plant pre-miRNAs from miRBase version 10.0; Griffiths-Jones et al., 2006) and a set of true non-miRNA sequences (tRNAs, 5S rRNA, 5.8S rRNA, and mRNAs; in total about 20,000 sequences). The filter steps discard non-miRNA sequences and favor pre-miRNAs; the filters are as follows (Zhang et al., 2006; Ritchie et al., 2007): subsequences of length ≤400 nt with any nucleotide fraction below 0.1 are discarded; subsequences forming structures [predicted for sequence length ≤400 nt by RNASHAPES (Steffen et al., 2006) with ∆G_{calc} = 0.1 kcal/mol or for longer sequences by RNALFold (Hofacker et al., 2004) with a maximal separation of base pairs of 400 nt at 37°C] with a normalized energy ∆G/length(structure)/f_{GC} ≤−0.75 kcal/mol/nt are retained: subsequences forming stem-loops shorter than 30 characters in the alignment-like format, possessing no helix with at least 7 consecutive base pairs, folding into structures with too many junctions or loops [length(sequence)/max length(stem-loops) ≥6], or not possessing structural elements with at least a fraction of 0.75 base pairs in a sequence window of 20 characters are discarded.

If a sequence (and structure) remains after the filter steps, YAMP decides on its possibility to be a pre-miRNA using a Hidden-Markov model with four transition states (is-miRNA, is-non-miRNA, is-non-miRNA→is-miRNA, is-non-miRNA). In case of a positive decision, YAMP searches in the stem-loop parts for a miRNA/miRNA* duplex in a window of 20 characters in the alignment-like format with a Hidden-Markov model. Transition probabilities are learned from true miRNA/miRNA* duplexes from A. thaliana. YAMP exhibits high sensitivity (~85%) and specificity (99.8%), which was demonstrated in several validation steps (for example with non-miRNAs from Rfam (Griffiths-Jones et al., 2005), pseudo-hairpins from Homo sapiens (Ng and Mishra, 2007), and pseudo-hairpins from A. thaliana) and was supported by a genomic search for miRNAs in the A. thaliana genome.

RNA extraction and semi-quantitative RT PCR

For semiquantitative RT PCR, total RNA was isolated from 100 mg of tomato leaf tissue using CONCERT (Plant RNA Purification Reagent, Invitrogen), followed by RNA purification and a DNA cleavage step on columns of RNeasy Plant Total RNA kit (Qiagen). RNA samples (2 µg) were applied to RT PCR reactions with 5’ ToTCP (5’GAAAAGCCGCCACAGC3’) and 3’ ToTCP (5’GCCTTTTTGATAAGCCAATC3’) primers derived from conserved motifs of tomato TCP gene (AC EF091571). RT PCR reactions were performed by Titan One Tube RT PCR System (Roche). A reaction mixture was used according to the manufacturer’s manual except that concentration of primers
was increased to 1.5 μM. Reverse transcription was run for 15 min at 50°C; after 2 min denaturation at 94°C, the polymerase chain reaction was started with cycles of 20 s at 94°C, 20 s at 54°C and 30 s at 68°C. Aliquots of the PCR product (10 μl) were taken at annealing steps of the 25th, 30th, 35th cycle and, together with the end product after the 38th cycle, were electrophoresed in 2% agarose gels; after staining with ethidium bromide, the fluorescence signal was quantified using the TYPHOON Phospholmager (Amersham Biosciences, Sunnyvale, California, USA) using ImageQuant software (Molecular Dynamics, Sunnyvale, California, USA).

Acknowledgments

The project was supported by a grant from the German Research Foundation to DR and GS and from ASCR I055015S58 to JM.

References


Figure S1. Northern blot analysis of vsRNAs in a total RNA extract from PSTVd QFA- and AS1-infected tomato plants. This figure is identical to Fig. 3 except that exposition time of all Northern blots was 1 day.

Figure S2. Graphical overview of probes used for Northern hybridization. Part A shows probes in (+) orientation, part B in (−) orientation.
Table S1. PCR primers for synthesis of PSTVd probes. Nucleotide positions of (+)-stranded PSTVd are given. A T7 promoter sequence (5’-TAA TAC GAC TCA CTA TA-3’) is abbreviated by T7; in a few cases the promoter sequence is followed by additional (PSTVd unrelated) Gs to increase efficiency of *in vitro* transcription. Primers were synthesized by MWG-Biotech, Germany, and biomers.net, Germany. Resulting PSTVd probes are depicted graphically in Fig. S2; start and end positions of probes are listed in Table 1 left.

<table>
<thead>
<tr>
<th>Name</th>
<th>Orientation</th>
<th>Position</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>For probes with (−) orientation:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF21</td>
<td>(+)</td>
<td>1– 20</td>
<td>GGG AAC TAA ACT CTT GGT TC</td>
</tr>
<tr>
<td>ND1</td>
<td>(+)</td>
<td>61– 78</td>
<td>GAA GGC GCC TCG GAG GAG</td>
</tr>
<tr>
<td>AF9</td>
<td>(+)</td>
<td>180– 207</td>
<td>TCA OCC TTC CTT TCT TCG GGT GTC CTT C</td>
</tr>
<tr>
<td>AF6</td>
<td>(+)</td>
<td>268– 292</td>
<td>GGA AAC AAG TGA AGC TCC CCA GAA C</td>
</tr>
<tr>
<td>RGV1</td>
<td>(−)</td>
<td>87– 69</td>
<td>T7- GG- CCT GAA GCG CTC CTC CGA G</td>
</tr>
<tr>
<td>ND2</td>
<td>(−)</td>
<td>208– 181</td>
<td>T7- GGA AGG ACA CCC GAA GAA AAG GAT G</td>
</tr>
<tr>
<td>XC3</td>
<td>(−)</td>
<td>295– 268</td>
<td>T7− G − GCG GTT CTC GGG ACG TTC ACT GCT TGC TTC</td>
</tr>
<tr>
<td>ND3</td>
<td>(−)</td>
<td>21– 1</td>
<td>T7− GGA ACC ACC AGT TTA GTT CCG</td>
</tr>
<tr>
<td>For probes with (+) orientation:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS2</td>
<td>(+)</td>
<td>353–359/1–14</td>
<td>T7− GGT TCC TCG GAA CTA AAC CCG</td>
</tr>
<tr>
<td>PS3</td>
<td>(+)</td>
<td>67– 87</td>
<td>T7− GGC TCG GAG GAG CCG TTC AAG</td>
</tr>
<tr>
<td>PS8</td>
<td>(+)</td>
<td>175– 197</td>
<td>T7− GGT TTT CAC CCT TCC TTT CCT C</td>
</tr>
<tr>
<td>PS6</td>
<td>(+)</td>
<td>249– 266</td>
<td>T7− G − GCT TCG GGT ACT AOC CCG</td>
</tr>
<tr>
<td>PS1</td>
<td>(−)</td>
<td>91– 115</td>
<td>CAG TTC GGT CCA GGT TCC CCC GGG G</td>
</tr>
<tr>
<td>PS4</td>
<td>(−)</td>
<td>153– 181</td>
<td>GAA AAC CCT GTT TCG GCG GGA ATT ACT CC</td>
</tr>
<tr>
<td>PS7</td>
<td>(−)</td>
<td>268– 295</td>
<td>GCG GTT CTC GGG ACG TTC ACT GCT TGC TTC C</td>
</tr>
<tr>
<td>PS5</td>
<td>(−)</td>
<td>337–359/ 1–2</td>
<td>GGA GGA ACC AAC TGC GGT TCC AAG G</td>
</tr>
</tbody>
</table>