UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Doctoral program in Biotechnology

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

INSTITUTO DE BIOLOGÍA MOLECULAR Y CELULAR DE PLANTAS (IBMCP)

INSTITUTO DE BIOLOGÍA INTEGRATIVA DE SISTEMAS (I^2SysBIO)

DOCTORAL THESIS

Novel insights of viroid biology and host responses to their infection

Presented by:

Joan Márquez Molins

Directors:

Dr. Gustavo Germán Gómez

Dr. José Antonio Navarro Bohígues

Dr. Vicente Pallás Benet

TO OPT TO THE DEGREE OF DOCTOR BY THE
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Valencia, March 2022

ACKNOWLEDGEMENTS

Arribar fins a aquest moment no hauria sigut possible sense l'inestimable suport de les persones que m'han rodejat. En primer lloc, he d'agrair a la meua excepcional família que ha estat al meu costat durant tot el camí. Ma mare ha sigut el meu suport incondicional i no tinc suficient paraules per agrair-li-ho. Tot i els moments de crisi, els daltabaixos emocionals, sempre he estat rodejat de bona gent, d'amics i companys extraordinaris. Aquesta experiència a cabal de dos laboratoris, de dos instituts, de dos universitats, m'ha permet a més d'aprendre, conèixer persones meravelloses. He gaudit amb els experiments, he aprés a escriure en anglès i a desenvolupar les meues pròpies idees. També m'he adonat que organitzar no és fàcil, però que m'agrada, i que el més important és la empatia. Aquesta tesi l'he fet per l'entusiasme de Gustavo, l'excel·lència de Toni i la saviesa de Vicente. En especial estic molt agraït a Toni perquè es qui m'ha ensenyat a treballar al laboratori. També estic molt agraït a Ricardo Flores, per les seues ensenyances sobre els viroids a les classes de màster. La seua incansable dedicació va despertar la meua curiositat sobre aquests àcids nucleics tan particulars, i sempre serà un referent. A més, haver compartit moments amb científics amb tanta experiència com Mari Carmen Marqués, Jesús, Fede, Mari Carmen Herranz, Mikhail, Techu, Mireya i Antonio ha sigut de gran ajuda, amb una menció especial per a Lorena que gestiona el laboratori de meravella i que m'ha ajudat molt. També estic molt agraït amb els bioinformàtics amb els que he coincidit: Alejandro Sanz, Luis Cervera, Luis Orduña i a l'última etapa Pascual i Julia, així com amb els futurs doctors experimentals María Saiz, David, Luis i Gabriela. A l'estada a Suècia, estic molt agraït perquè Germán no em podia haver acollit millor, va ser una experiència molt enriquidora. Tampoc m'oblide dels estudiants als que he intentat ensenyar: Fran, Irene, Isi, Naiara, Inés, Laura, María Urrutia, ni dels companys de l'I2SysBio (EvolSysVir i TomsBio Lab). No obstant, al científic que més li he d'agrair és al Pinche, qui s'ha convertit en un gran amic i m'ha donat ànims quan més els necessitava. És una persona d'una sensibilitat extraordinària de la que em sent afortunat d'haver conegut. I es que per molt meravellosa que siga la ciència, no és el més important, perquè el que realment ens porta la felicitat són les persones.

La Conselleria d'Educació, Investigació, Cultura i Esports (Generalitat Valenciana) y el Fondo Social Europeo (FSECV 2014-2020) han cofinanciado la contratación del doctornado como personal investigador de carácter predoctoral (ACIF/2017/114) y una estancias predoctorales fuera de la Comunitat Valenciana (BEFPI/2020). La realización de esta tesis doctoral también se ha realizado en el marco de dos proyectos de investigación del Ministerio de Ciencia, Innovación y Universidades, con cofinanciación de fondos FEDER [BIO2017-88321-R y AGL2016-79825-R] .

INDEX

ABBREVIATIONS

AGO: *Argonaute protein.*

ASBVd: *Avocado sunblotch viroid.*

ASSVd: *Apple scar skin viroid.*

CCCVd: *Coconut cadang cadang viroid.*

CChMVd: *Chrysanthemum chlorotic mottle viroid.*

CCR: *central conserved region.*

cDNA: *complementary DNA.*

CEVd: *Citrus exocortis viroid.*

cTP: *chloroplast transit peptide.*

circRNAs: *circular RNAs*.

CLVd: *Columnea latent viroid.*

CSVd: *Chrysanthemum stunt viroid.*

DCL: *Dicer-like; Ribonuclease III-like.*

DNA: *deoxyribonucleic acid.*

DRM: *domains rearranged methyltransferase.*

ELVd: *Eggplant latent viroid.*

en-circRNAs: *endogenous circular RNAs.*

ex-circRNAs: *exogenous circular RNAs.*

GFP: *green fluorescent protein.*

hc-siRNAs: *heterochromatic small interfering RNAs.*

HDA6: *histone deacetylase 6.*

HSVd: *Hop stunt viroid.*

miRNA: *microRNA.*

mCherry: *monomeric cherry fluorescent protein.*

mRNA: *messenger RNA.*

ncRNA: *non-coding RNA.*

NGS: *next generation sequencing.*

nt: *nucleotide.*

ORF: *open reading frame.*

P: *pathogenic domain.*

PAGE: *polyacrylamide gel electrophoresis.*

PCR: *polymerase chain reaction.*

PLMVd: *Peach latent mosaic viroid.*

PSTVd: *Potato spindle tuber viroid.*

PTGS: *post-transcriptional gene silencing.*

rDNA: *ribosomal DNA.*

RdDM: *RNA-directed DNA methylation.*

RDR: *RNA-dependent RNA polymerase.*

RISC: *RNA-induced silencing complex.*

RNA: *ribonucleic acid.*

RNA-pol II: *RNA polymerase II*.

rRNA: *ribosomal RNA.*

siRNA: *small interfering RNA.*

sRNA: *small RNA.*

ssRNA: *single-stranded RNA.*

ta-siRNA or tasiRNA: *trans-acting small interfering RNA.*

TBE: *Tris/borate/EDTA*.

TCH: *terminal conserved hairpin.*

TCR: *terminal conserved region.*

TIC: *translocon on the inner chloroplast membrane*

TL: *left terminal domain.*

TOC: *translocon on the outer chloroplast membrane*

TR: *right terminal domain.*

UTR: *untranslated region.*

V: *variable domain.*

vd-sRNAs: *viroid small RNAs.*

WGBS: *whole genome bisulfite sequencing.*

YFP: *yellow fluorescent protein.*

ABSTRACT

Viroids are the simplest pathogens with autonomous replication and have only been found naturally infecting higher plants. Since viroids were discovered in the seventies, we have gained considerable knowledge about their nature and replication mechanisms in host plants. However, many aspects of viroid biology are yet to be discovered. Therefore, a deeper understanding of the nature and mode of action of viroids have been the encompassing main goals of this thesis. For this purpose, simple and efficient procedures for obtaining infectious cDNA clones are essential. A new efficient method for constructing infectious viroid clones was developed and tested with one viroid of each family: eggplant latent viroid (ELVd, *Avsunviroidae*) and hop stunt viroid (HSVd, *Pospiviroidae*). This procedure was based on type IIS restrictions enzymes that cut outside of the recognition site and supposes a universal procedure for obtaining infectious clones of a viroid independently of its sequence, with a high efficiency.

Despite viroids have been considered as plant-pathogenic non-coding RNAs since their discovery, our computational analysis predicted small open reading frames in each of the HSVd and ELVd genomes. No significant similarities with proteins in the database of higher plants were found, but some of these predicted peptides were highly conserved among all HSVd and ELVd variants. Interestingly, the fusion of these conserved sequences to a fluorescent protein revealed a specific subcellular localization in the corresponding organelle where replication/accumulation takes place for each viroid: nucleolus and chloroplast for HSVd and ELVd, respectively. Mutations that truncate the nucleolar domain of HSVd were detrimental for the viroid while truncating any of the two ELVd ORF that contains a chloroplast transit signal also diminished (but to a lesser extent) viroid biological efficiency, maybe because of functional redundancy. Circular forms of both, HSVd and ELVd RNAs were found in polysome fractions, revealing their physical interaction with the translational machinery of the plant cell. Altogether, these experimental observations indicate that the coding capacity of viroids cannot be ruled out, although the definitive evidence (detection of the circRNA-encoded peptides) is a technological challenge to be addressed in future research lines.

Finally, to study the host changes that are produced during a symptomatic viroid infection, an integrative analysis of the timing and intensity of the genome-wide alterations in cucumber plants infected with HSVd was performed. Differential host transcriptome, sRNAnome and methylome were integrated to determine the temporal response to viroid-infection. Our results support that HSVd promotes the redesign of the cucumber regulatory-pathways predominantly affecting specific regulatory layers at different infection-phases. The initial response was characterized by a reconfiguration of the host-transcriptome by differential exon usage, followed by a predominant

down-regulation of the transcriptional activity modulated by the host epigenetic changes associated to infection and characterized by increased hypermethylation. The alterations in host sRNA and microRNA metabolism were marginal and mainly occurred at the late stage. Overall, these data constitute the first comprehensive map of the plant responses to a viroid infection.

RESUMEN

Los viroides son los patógenos con replicación autónoma más simples y sólo se han encontrado de forma natural infectando plantas superiores. Desde que se descubrieron en los años setenta, se ha adquirido un conocimiento considerable sobre su naturaleza y mecanismos de replicación en las plantas huésped. Sin embargo, aún quedan por descubrir muchos aspectos de la biología de los viroides. Por lo tanto, un conocimiento más profundo de la naturaleza y el modo de acción de los viroides han sido los objetivos principales que engloban esta tesis. Para ello, es esencial contar con procedimientos sencillos y eficientes para la obtención de clones de ADNc infecciosos. Se desarrolló un nuevo método eficiente para construir clones de viroides infecciosos y se probó con un viroide de cada familia: El viroide latente de la berenjena (ELVd, *Avsunviroidae*) y el viroide del lúpulo (HSVd, *Pospiviroidae*). Esta aproximación se basó en enzimas de restricción de tipo IIS que cortan fuera del sitio de reconocimiento y supone un procedimiento universal para obtener clones infecciosos de un viroide independientemente de su secuencia, con una alta eficiencia.

A pesar de que los viroides han sido considerados como ARN no codificantes desde su descubrimiento, nuestro análisis computacional predijo pequeños marcos de lectura abiertos en cada uno de los genomas de HSVd y ELVd. No se encontraron similitudes significativas con las proteínas de la base de datos de plantas superiores, pero algunos de estos péptidos predichos estaban altamente conservados entre todas las variantes de HSVd y ELVd. Curiosamente, la fusión de estas secuencias conservadas con una proteína fluorescente reveló una localización subcelular específica en el correspondiente orgánulo donde tiene lugar la replicación/acumulación para cada viroide: nucleolo y cloroplasto para HSVd y ELVd, respectivamente. Las mutaciones que truncan el dominio nucleolar de HSVd fueron perjudiciales para el viroide, mientras que el truncamiento de cualquiera de los dos ORF de ELVd que contiene una señal de localización al cloroplasto también disminuyó (pero en menor medida) la eficiencia biológica del viroide, tal vez debido a la redundancia funcional. Se encontraron formas circulares de los ARN de HSVd y ELVd en fracciones polisómicas, lo que revela su interacción física con la maquinaria de traducción de la célula vegetal. En conjunto, estas observaciones experimentales indican que no se puede descartar la capacidad de codificación de los viroides, aunque la prueba definitiva (la detección de los péptidos codificados por los circRNAs) es un reto tecnológico que deberá abordarse en futuras líneas de investigación.

Finalmente, para estudiar qué cambios se producen en el huésped durante la infección con un viroide sintomático, se realizó un análisis integrador de las alteraciones genómicas de plantas de pepino infectadas con HSVd. Se integraron los transcriptomas, el sRNAnomas y el metilomas para determinar la respuesta temporal a la infección por el viroide. Nuestros resultados apoyan que el

HSVd promueve el rediseño de las vías reguladoras del pepino afectando predominantemente a capas reguladoras específicas en diferentes fases de la infección. La respuesta inicial se caracterizó por una reconfiguración del transcriptoma del hospedador mediante el uso diferencial de exones, seguido de una predominante regulación a la baja de la actividad transcripcional modulada por los cambios epigenéticos del hospedador asociados a la infección y caracterizada por un aumento de la hipermetilación. Las alteraciones en el metabolismo de los ARN pequeños y microARNs del huésped fueron marginales y se produjeron principalmente en la fase tardía. En general, estos datos constituyen el primer mapa exhaustivo de las respuestas de la planta a la infección de un viroide.

RESUM

Els viroids són els patògenes més simples amb replicació autònoma i només s'han identificat de forma natural infectant a plantes superiors. Des que es descobriren als anys setanta, s'ha adquirit un coneixement considerable sobre la seua natura i els mecanismes de replicació en plantes hoste. No obstant, encara queden per descobrir molts aspectes de la biologia dels viroids. Per tant, un coneixement més profund de la natura i el mode d'acció dels viroids han sigut els objectius principals que engloben aquesta tesi. Per a això, és essencial la disponibilitat de procediments senzills i eficients per a l'obtenció de clones infecciosos. Es va desenvolupar un nou mètode eficient per a construir clones infecciosos y es fa provar amb un viroid de cada família: el viroide latent de la albergínia (ELVd, *Avsunviroidae*) y el viroid del llúpol (HSVd, *Pospiviroidae*). Aquesta aproximació es basà en enzims de restricció de tipus IIS que tallen fora del lloc de reconeixement i suposa un procediment universal per obtenir clones infecciosos de un viroid independentment de la seua seqüencia amb una elevada eficiència.

Tot i que els viroids s'han considerat com ARNs no codificants des del seu descobriment, el nostre anàlisi computacional va predir xicotets ORF als genomes de HSVd y ELVd. No es trobaren similituds significatives amb proteïnes depositades a les bases de dades, però alguns d'aquest pèptids estaven altament conservats a les variants de HSVd y ELVd. Curiosament, la fusió d'aquestes seqüencies conservades amb una proteïna fluorescent revelà una localització subcel·lular especifica al orgànul on te lloc la replicació/acumulació de cada viroid: nuclèol i cloroplast per a HSVd i ELVd, respectivament. Les mutacions que trunquen el domini nucleolar de HSVd foren perjudicials per al viroid, mentre que el truncament de qualsevol de les dos ORF de ELVd que contenen una senyal de localització al cloroplast també va disminuir (però en menor mesura) l'eficiència biològica del viroid, el que pot ser degut a una redundància funcional. Es detectaren formes d'ARN circular de HSVd i ELVd a les fraccions polisòmiques, el que revela la seua interacció física amb la maquinaria de traducció cel·lular. En conjunt, aquestes observacions experimentals indiquen que no es pot descartar la capacitat codificants dels viroids, encara que la evidencia definitiva (la detecció del pèptids codificats per ARN circulars) es un repte tecnològic que s'haurà d'adreçar en línies d'investigació futures.

Finalment, per tal d'estudiar que canvis es produeixen a l'hoste durant la infecció amb un viroid simptomàtic, es va realitzar un anàlisi integrador de les alteracions genòmiques de les plantes de cogombre infectades amb HSVd. S'integraren els transcriptomes, sARNomes i metilomes per determinar la resposta temporal a la infecció per viroid. Els resultats obtinguts suporten que HSVd promou un redisseny de les vies reguladores de cogombre afectant predominantment a nivells reguladors específics a les diferents etapes de la infecció. La resposta inicial es caracteritzà per una

reconfiguració del transcriptoma de l'hoste mitjançant l'ús diferencial d'exons, seguit d'una repressió transcripticional modulada per canvis epigenètics de l'hoste caracteritzats per una major hipermetilació. Les alteracions al metabolisme de ARN xicotets i microARNs de l'hoste van ser marginals i es produïren principalment al final de la infecció. En general, aquestes dades constitueixen el primer mapa exhaustiu de les respostes de la planta a la infecció per un viroid.

INTRODUCTION

1. General characteristics of viroids

Viroids are the simplest pathogens with autonomous replication currently known and have only been found naturally infecting higher plants (Flores et al., 2015). The mature form of viroids is a covalently closed single-stranded small RNA molecule (246-434 nt) with a compact structure due to a high degree of self-complementarity (Flores et al., 2004; Navarro et al., 2021a). Whether viroids may be coding or not was firstly investigated in several cell-free protein-synthesizing systems, which resulted in two independent publications in 1974 (Davies et al., 1974; Hall et al., 1974). No specific proteins were detected in these experiments, and the amino acid incorporation was not stimulated much above the control without the viroid RNA, being only a tiny fraction when compared with that observed with viral RNAs. Those results together with the small size (the RNA encoding, if any, translation products of minimal complexity) and the circularity observed by electron microscopy (Sanger et al., 1976), which hampered translation by conventional ribosome scanning, led to the paradigm of viroids as non-coding RNAs. This notion was reinforced by the sequencing of viroid RNAs, which confirmed the circular nature and the absence of canonical open reading frames (Gross et al., 1978).

Viroids replicate through a rolling circle mechanism with RNA intermediates (Branch and Robertson, 1984; Wang, 2021) and subvert host factors for completing their life cycle (Adkar-Purushothama and Perreault, 2020). In infected tissues, both polarities of the viroid genome are present, and the plus polarity (+) is arbitrarily assigned to the most abundant strand. These pathogens are the causal agents of several plant diseases showing similar symptomatology to those associated with some plant viruses (Owens and Hammond, 2009; Navarro et al., 2012a; Flores et al., 2016). However, in many cases, viroid infection is asymptomatic or causes mild symptoms (Gómez et al., 2008; Daròs, 2016). Viroids spread mainly by vegetative propagation of infected plants or by mechanical transmission by pruning tools, although seed transmission has also been reported (Matsushita et al., 2018). Transmission by bumblebees and whiteflies has been described, but the relevance of these transmission pathways in nature remains unclear (Škorić, 2017). According to their structural and functional features, viroids are classified into two families: *Pospiviroidae* and *Avsunviroidae* (Di Serio et al., 2014) (Figure I.1). Family *Pospiviroidae* is comprised by five genera (*Pospiviroid, Hostuviroid, Cocadviroid, Apscaviroid and Coleviroid*) while the family *Avsunviroidae* has three genera (*Avsunviroid, Elaviroid and Pelamoviroid*).

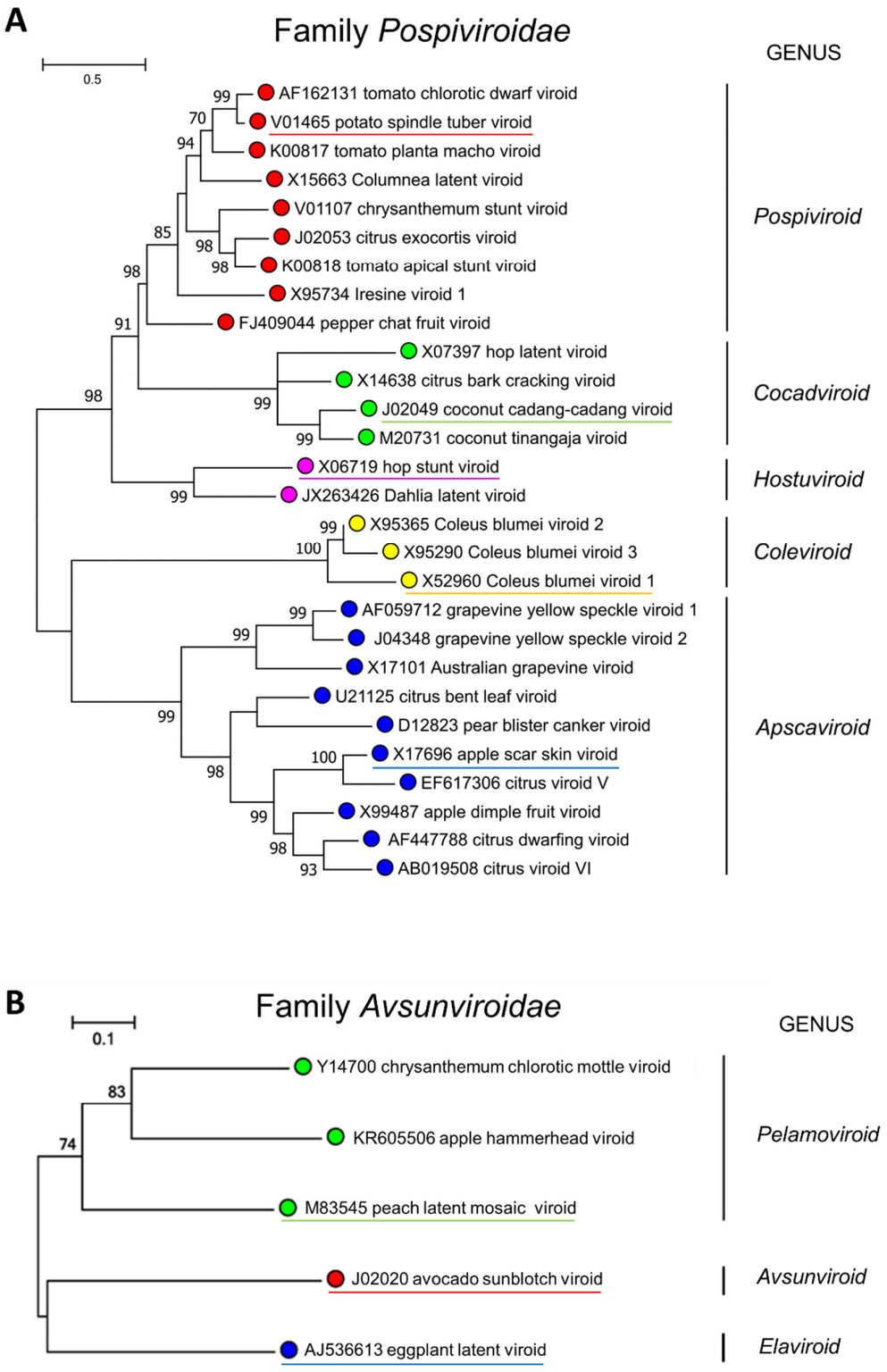


Figure I.1. Phylogenetic trees for the viroid families *Pospiviroidae* (A) and *Avsunviroidae* (B). Viroids cluster according to their respective taxonomic position within the recognized genera (denoted on the right). The type species of each genus is underlined. Trees were inferred by the neighbor-joining method (1000 of bootstrap value) and evolutionary distances were estimated according to the model of Jukes and Cantor. Bootstrap values are reported at the nodes. The phylogenetic trees are reproduced from the ICTV Taxonomy Profiles (Di Serio et al., 2018; Di Serio et al., 2021).

1.1 Family *Pospiviroidae*

Most currently known viroids belong to the family *Pospiviroidae* (Di Serio et al., 2021). They are characterized by having a rod-like secondary structure with a central conserved region (CCR), nuclear replication through an asymmetrical rolling circle mechanism and generally having a broad host range. The CCR is formed by two sections of conserved nucleotides located in the upper and lower strands flanked by inverted repeats and it is the main structural feature adopted for the demarcation of the five genera of the family (Figure I.2). Replication of *Pospiviroidae* members is possible by the subversion of the host DNA-dependent RNA polymerase II (RNA pol II) to transcribe multimeric RNAs from the circular mature form (Mühlbach and Sänger, 1979; Spiesmacher et al., 1983; Hutchins et al., 1985; Wang, 2021). The cleavage of these intermediates is carried on by a host type-III RNase (Gas et al., 2007) and the resultant monomers are circularized by a host DNA ligase 1 (Nohales et al., 2012a).

The intra-nuclear localization has been deeply investigated for the potato spindle tuber viroid (PSTVd) (Harders et al., 1989; Qi and Ding, 2003), the type species of this family and the first discovered viroid (Diener, 2003). *In situ* hybridization studies with specific probes revealed that the (−) strand RNA was localized in the nucleoplasm, whereas the (+) strand RNA was additionally localized in the nucleolus (Qi and Ding, 2003). This differential sub-nuclear localization pattern implies the existence of a selective mechanism of trafficking for the (+) strand RNA (Qi and Ding, 2003). Additionally, this nuclear compartmentalization has been reported in other pospiviroids (Bonfiglioli et al., 1996), suggesting that it is probably a common feature of all the members of the family *Pospiviroidae*. Even though some viroids of this family have been reported to exclusively infect a single host (Verhoeven et al., 2013), many others affect dozens of plants of different families, including crops of economic importance (Hammond, 2017).

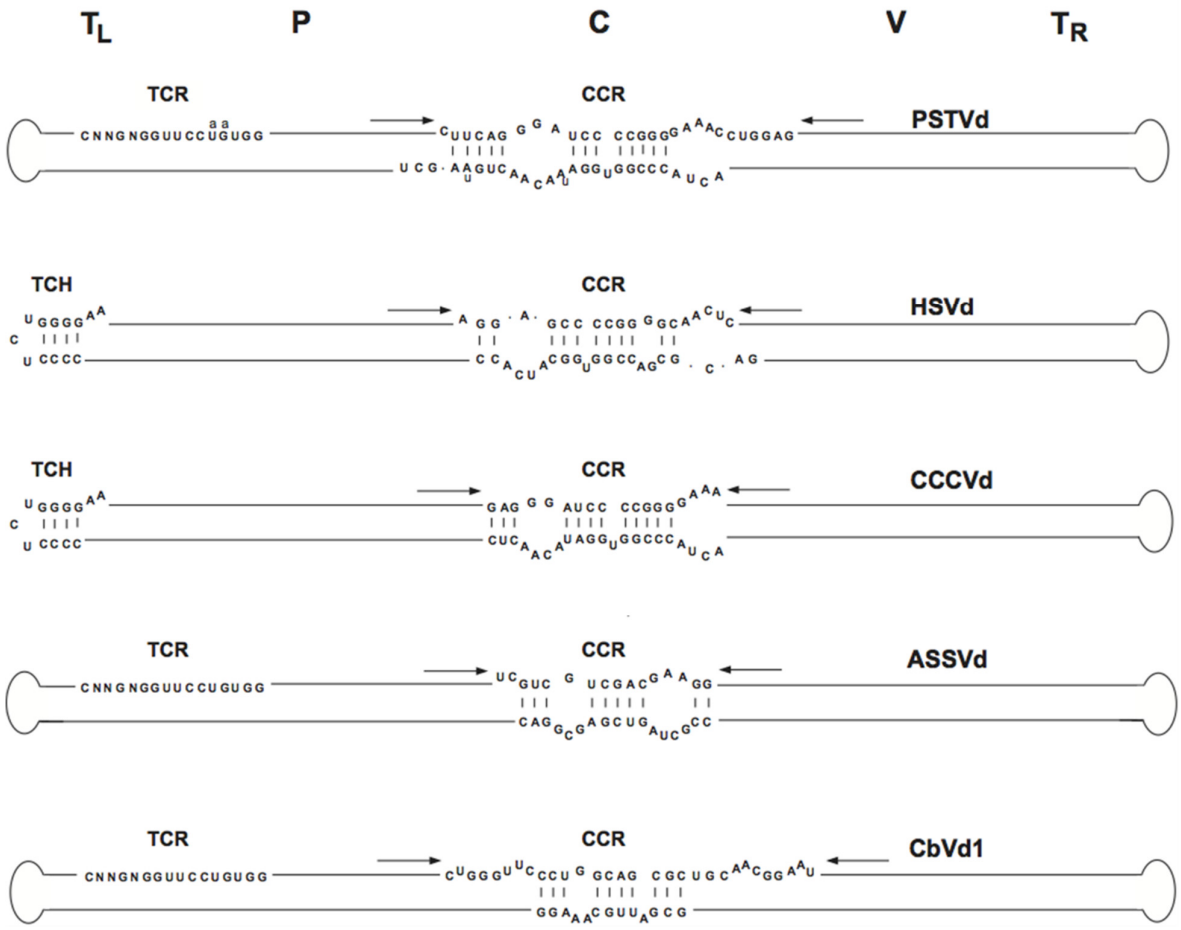


Figure I.2. Schematic structures of viroids in the five genera of the family *Pospiviroidae*. The type species of each genus is indicated on the right together with the approximate location of the five structural domains shown above. Core sequences of the central conserved region (CCR), terminal conserved region (TCR) and terminal conserved hairpin (TCH) are shown. Arrows indicate flanking sequences, which, together with the core nucleotides of the CCR upper strand, form imperfect inverted repeats. The approximate location of the five domains, C (central), P (pathogenic), V (variable) and TL and TR (terminal left and right), at the top of the figure. Figure reproduced from the ICTV Report Virus Taxonomy Profile: *Pospiviroidae* (Di Serio et al., 2021)

The type species of the family *Pospiviroidae* is *Potato spindle tuber viroid* because it was the first viroid discovered (Diener, 1971). However, in this thesis, *Hop stunt viroid* has been chosen as a representative member because of the considerable amount of experiments that had been conducted with this viroid in our research group, shedding light on several aspects of host-viroid interactions (Marquez-Molins et al., 2021).

1.2 Family *Avsunviroidae*

On the other hand, the members of the family *Avsunviroidae* are comprised of RNA molecules either with a quasi-rod-like or a branched secondary structure, that exhibit plastid replication (mostly in chloroplasts) via a symmetrical rolling mechanism (Di Serio et al., 2014). These self-replicating circular RNAs lack a CCR and display less extensive internal base-pairing and more structural diversity than members of the family *Pospiviroidae* (Figure I.3A). Nevertheless, there is a unifying structural hallmark in the family constituted by the existence in both polarity strands of autocatalytic hammerhead ribozymes active at self-cleaving the oligomeric RNAs (Hutchins et al., 1986) (Figure I.3B). Additionally, guanine + cytosine content and solubility in 2 M LiCl, are used to allocate the four recognized species among three different genera (Figure I.1B). All these viroids show an extremely narrow host range, which is restricted to the plant in which they were discovered or to very closely related species of the same genus: avocado (Symons, 1981), peach (Flores et al., 2006), chrysanthemum (Navarro and Flores, 1997), eggplant (Fadda et al., 2003) or apple (Serra et al., 2018). In contrast to the nuclear compartmentalization of *Pospiviroidae* members, members of the family *Avsunviroidae* accumulate and as mentioned above, replicate in the plastids. Interestingly, they display the highest mutation rate reported for any biological entity (Gago et al., 2009). This is probably related to the very low fidelity of the chloroplastic RNA polymerase involved in their replication process (Navarro et al., 2000). Moreover, a chloroplastic isoform of tRNA ligase is involved in the ligation of the self-cleaved monomers (Nohales et al., 2012b).

Even though the type species of this family is *Avocado sun blotch viroid,* the difficulties of performing experiments in avocado, a woody tree that it is the only host of this viroid, we have opted for using *Eggplant latent viroid*. This viroid has been proposed as a friendly experimental system in the family (Daròs, 2016) and eggplant is an herbaceous crop with economic importance.

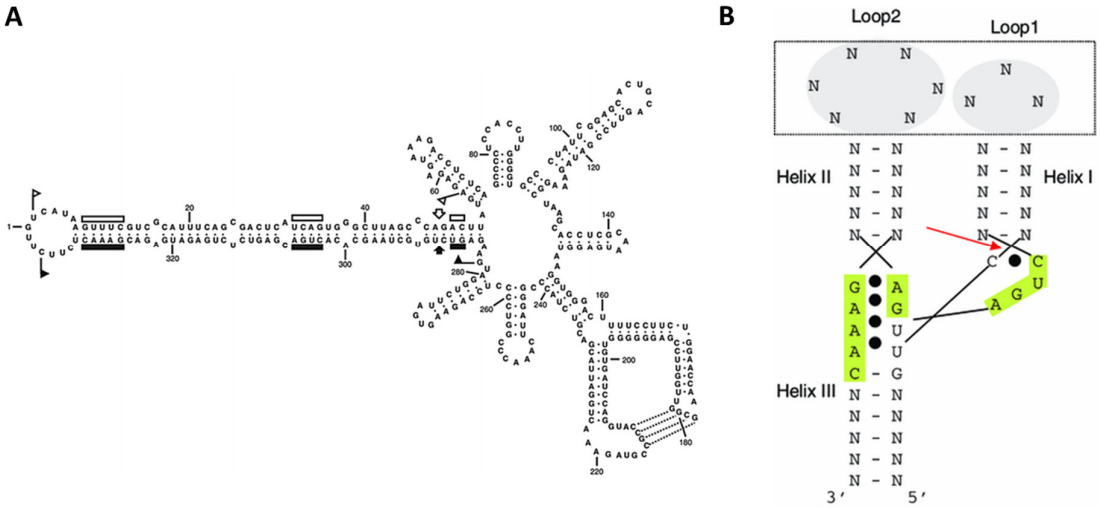


Figure I.3. Distinctive features of the *Avsunviroidae* family A) Secondary structure of peach latent mosaic viroid (PLMVd) a representative member of the *Avsunviroidae* family. Plus and minus self-cleavage domains in PLMVd genomic and anti-genomic RNAs are delimited by flags, the eleven conserved residues present in most natural hammerhead structures are indicated by bars, and the self-cleavage sites by arrows. Solid and open symbols refer to plus and minus polarities, respectively. A kissing-loop interaction between positions 178-181 and 211-214 is shown **B**) Schematic representation of the structure of a hammerhead ribozyme. The arrowhead marks the self-cleavage site, and the tertiary interaction between loops 1 and 2 (in grey), which facilitates the catalytic activity of the ribozyme at the low magnesium concentration existing *in vivo*, is indicated by a dotted rectangle. Continuous lines and dots between nucleotides denote canonical and non-canonical base pairs, respectively. Sequences strictly or highly conserved in natural hammerhead structures are shown within green boxes. N, any nucleotide forming stable helices I, II and III. Reproduced from Di Serio *et al.* (2014).

2. *Hop stunt viroid*

Hop stunt viroid (HSVd) is a viroid of the family *Pospiviroidae* that infects both trees (e.g. *Prunus spp.* and *Citrus spp.*) and herbaceous plants (e.g. hop, grapes and cucumber) (Marquez-Molins et al., 2021). The main characteristics of this viroid are summarized in Table I.1. HSVd was first discovered in hop (*Humulus lupulus*) plants showing abnormal dwarfing of bines (Sasaki and Shikata, 1977) and since then has been identified as the causal agent of several diseases (Figure I.4). Interestingly, multiple studies using HSVd as a model have shed light on different aspects of viroid-host interactions (Figure I.5).

Table I.1 Main characteristics of HSVd

TAXONOMY:	*Hop stunt viroid* is the type species of the genus *Hostuviroid* (family *Pospiviroidae*). The other species of this genus is *Dahlia latent viroid*, which presents an identical central conserved region (CCR) but lacks other structural hallmarks present in HSVd. HSVd replication occurs in the nucleus through an asymmetric rolling-circle model as in the other members of the family *Pospiviroidae*, which also includes the genera *Pospiviroid, Cocadviroid, Apscaviroid,* and *Coleoviroid.*
PHYSICAL PROPERTIES:	HSVd consists of a single-stranded, circular RNA of 295–303 nucleotides depending on isolates and sequence variants. The most stable secondary structure is a rod-like or quasi-rod-like conformation with two characteristic domains: a CCR and a TCH similar to that of cocadviroids. HSVd lacks a TCR.
HOSTS and SYMPTOMS:	HSVd infects a very broad range of natural hosts and has been reported to be the causal agent of five different diseases (citrus cachexia, cucumber pale fruit, peach dapple, plum dapple and hop stunt) although in most hosts causes latent infections. It is distributed worldwide.
TRANSMISSION	HSVd is transmitted mechanically and by seed.

The ability of HSVd to infect cucurbitaceous hosts has been key for studying systemic movement, as it is relatively easy to obtain phloem exudate from these plants. At the beginning of this century, it was not known whether viroids move systemically as free RNAs or form ribonucleoprotein complexes, and, in general, plant mechanisms involved in long-distance RNA movement remained unclear (Ueki and Citovsky, 2001). However, in 2001, two independent research groups reported the interaction *in vitro* between HSVd and the phloem protein 2 (PP2), which is a dimeric lectin and the most abundant component of the cucumber phloem exudate (Gómez and Pallás, 2001; Owens et al., 2001). The formation of the ribonucleoprotein complex *in vivo* was confirmed by immunoprecipitation and further characterization by intergeneric graft assays revealed that the complex was translocated to the scion, which even developed symptoms of HSVd infection despite being a non-host plant (Gómez and Pallás, 2004). Moreover, it was reported that PP2 was the only phloem protein that binds HSVd in cucumber plants (Gómez and Pallás, 2004).

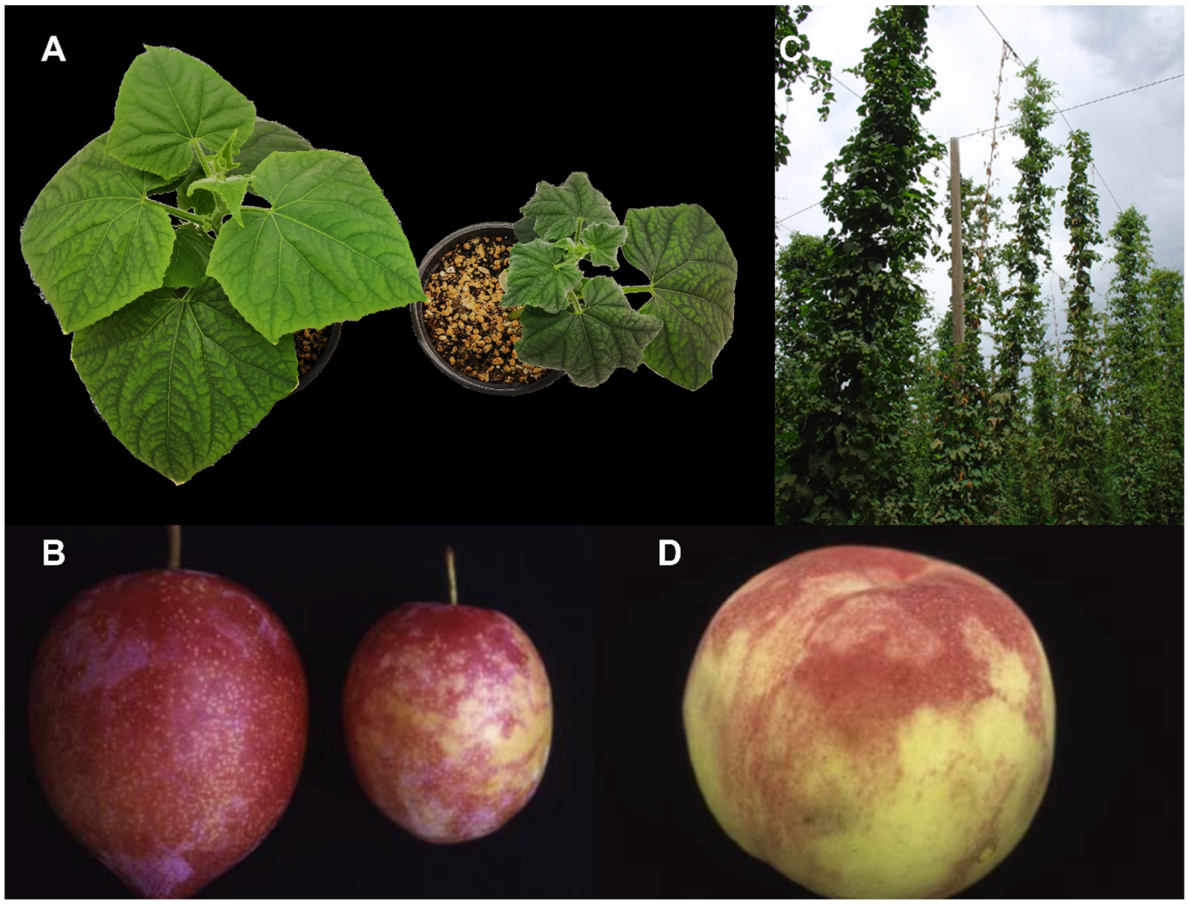


Figure I.4. Symptomatology associated with HSVd infection. A) Cucumber plant inoculated with mock inoculum (left) or HSVd (right) showing clear stunting and leaf distortion. **B**) Plum dapple fruit disease in plum cv. Shiho. Plum fruit with irregular reddish blotches on the pericarp caused by HSVd (right) compared with healthy (left) plum. **C**) Hop plants in field conditions showing abnormal dwarfing of bines. **D**) Peach dapple fruit disease in peach cv. Hakuhou. Chlorotic blotches on the pericarp of a fruit of peach infected with HSVd. Figure reproduced from Marquez-Molins *et al* (Marquez-Molins et al., 2021).

Nicotiana benthamiana is a frequent experimental model in the study of plant-pathogen interactions, however, it could not be infected by HSVd (either by the inoculation of nucleic acid preparations from viroid-infected plants or HSVd dimeric transcripts). In order to determine if it could be due to a defect in replication or systemic movement, transgenic plants of *N. benthamiana* expressing dimeric HSVd (+) cDNA (HSVd-Nb) were obtained (Gómez and Pallás, 2006). HSVd was correctly processed into circular forms and grafting assays showed that the movement to distal parts was possible. Therefore, deficiencies in the interaction with other host factors, related either to the early steps of replication or to cell-to-cell movement, may explain the reported lack of infectivity. Additionally, these HSVd-Nb plants were used to propose that mature HSVd forms are resistant to RNA silencing (Gómez and Pallás, 2007a), despite the extensive presence of viroid derived small RNAs (vd-sRNA) in infected tissues (Navarro et al., 2009; Martinez et al., 2010). Vd-sRNA in infected cucumber plants were mostly 21-, 22- and 24- nucleotides in length and derived equally from plus and minus HSVd RNA strands. Therefore, this result suggested that, at

least in nuclear replicating viroids, vd-sRNAs should be mainly originated from double-stranded RNA viroid replication intermediates (Gómez et al., 2009). Interestingly, phloem vd-sRNAs of cucumber accumulated preferentially as 22-nucleotide species with a consensus sequence over-represented, an observation that suggested the existence of a selective trafficking of vd-sRNAs to the phloem tissue (Martinez et al., 2010). Furthermore, the transgenic HSVd-Nb plants were analyzed at different growing temperatures and used as stocks in grafting assays with the rdr6i-Nb line, in which the RNA-dependent RNA polymerase 6 (RDR6) is constitutively silenced (Gómez et al., 2008). These results revealed a clear dependency of temperature and RDR6 activity for the apparition of phenotypical alterations upon HSVd infection (Gómez et al., 2008). Thus, it was concluded that symptom expression is dependent on an active state of the viroid-specific RNA silencing pathways and independent of mature HSVd accumulation levels.

Regarding strictly intracellular interactions of HSVd with host factors, the dependency of RNA pol II for replication of members in the *Pospiviroidae* family was first evidenced in a sequence variant of HSVd (initially identified as cucumber pale fruit viroid) (Mühlbach and Sänger, 1979). Previously, it was not known if an RNA or DNA-dependent polymerase would be responsible for the replication of these pathogens.

The interaction of HSVd with the histone deacetylase 6 (HDA6) was proposed to explain the methylation changes observed in the host genome (Martinez et al., 2014; Castellano et al., 2015; Castellano et al., 2016b). In infected cucumber plants, it was observed a high expression of ribosomal RNA (rRNA) precursors caused by the decrease of DNA methylation in the promoter region of rRNA genes during HSVd infection (Martinez et al., 2014). That study reported a previously unknown mechanism of epigenetic alteration associated with viroid infection in plants. A similar pattern of loss of cytosine methylation of usually silenced rRNA genes was detected in transgenic *N. benthamiana* plants constitutively expressing HSVd, which gave evidence that it might be a generalized phenomenon in HSVd pathogenesis (Castellano et al., 2015). The interaction between HSVd and HDA6 was detected *in vitro* and *in vivo*, while transient overexpression of recombinant HDA6 reverted the hypo-methylation status of rDNA, characteristic of HSVd-infected plants, and reduced viroid accumulation (Castellano et al., 2016b). Those observations led to the proposal that this viroid recruits and functionally subverts HDA6 to promote the host-epigenetic changes observed during viroid pathogenesis. Additionally, the transcriptional increase caused by the viroid-mediated HDA6 recruitment may favour a spurious recognition of HSVd-RNA as an RNA Pol II template, thereby improving viroid replication.

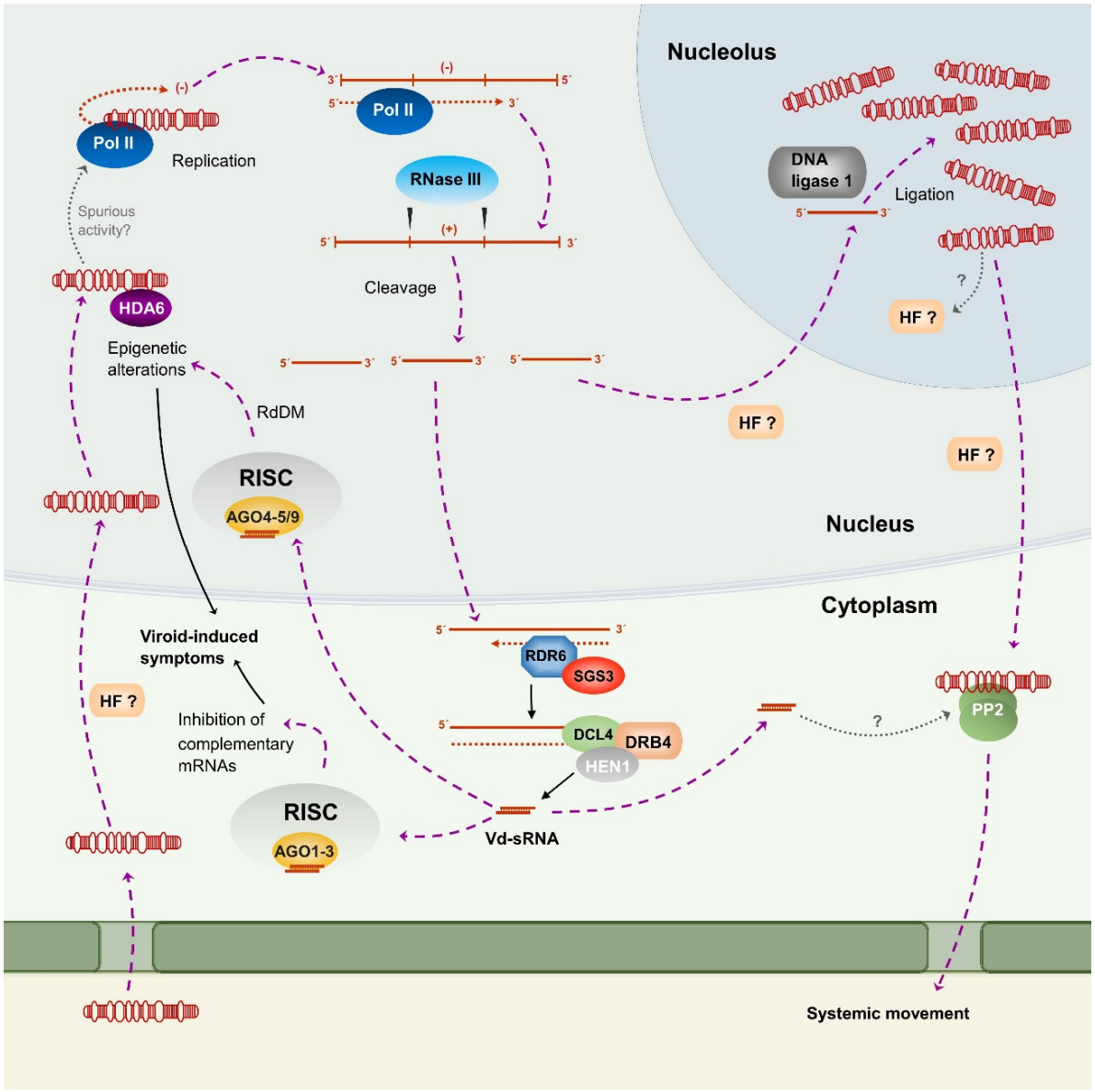


Figure I.5. Proposed model of HSVd cell cycle and interactions with host factors. HSVd RNA enters the cell through the plasmodesmata and it is trafficked to the nucleus, probably assisted by unidentified host factors. In the nucleus, HSVd subverts the activity of HDA6 and causes epigenetic alterations. The subversion of HDA6 may favour a spurious recognition of the genomic HSVd RNA by the DNA-dependent-RNA polymerase II (Pol II). HSVd replication follows asymmetric rolling circle in which Pol II transcribes HSVd RNAs. Oligomeric RNAs are cleaved by a host enzyme with RNase III activity and the resultant monomers are circularized by DNA ligase I, probably in the nucleolus, in which it is assumed that mature HSVd molecules accumulate. Other HSVd replication intermediates are recognized as aberrant transcripts and enter the trans-acting small interfering RNA (tasi-RNA) biogenesis pathway. Therefore, these RNAs are trafficked to the cytoplasm, transcribed by RDR6 and processed by DCL4 into vd-sRNAs which are loaded into different Argonaute proteins (AGO). Vd-sRNAs of 21-22- nucleotides are loaded in AGO 1-3 in the RISC complex and direct the inhibition of host mRNA transcripts in the cytoplasm. Moreover, vd-sRNAs of 24-nucleotides are primarily responsible for RNA-directed DNA methylation (RdDM) in the nucleus. Thus, HSVd-sRNAs cause transcriptional (nucleus) and posttranscriptional (cytoplasm) alterations which account for the viroid-induced symptoms. Mature HSVd RNAs are exported to the phloem by forming a ribonucleoprotein complex with the phloematic protein 2 (PP2). Additionally, the systemic movement of HSVd-sRNA may be also possible due to the interaction with PP2. Figure reproduced from Marquez-Molins *et al* (Marquez-Molins et al., 2021).

In addition to the alterations in vegetative tissues, HSVd infection is also associated with drastic changes in gametophyte development (Castellano et al., 2016a). It was observed that in pollen grain, the accumulation of HSVd induces a decondensation of the generative nucleus that correlates with a dynamic demethylation of repetitive regions in the host genome (Castellano et al., 2016a). Therefore, the authors proposed that HSVd infection impairs the epigenetic control of these repetitive regions, primarily rRNA genes and transposable elements, in gametic cells of cucumber. These observations revealed a previously undescribed phenomenon as a consequence of pathogen infection in this reproductive tissue.

3. *Eggplant latent viroid*

Eggplant latent viroid (ELVd) was first detected by Fagoaga *et al.* in a survey of viroid and viroid-like agents in a series of vegetable crop species (Fagoaga et al., 1994) and has been classified into a new genus named *Elaviroid* in the family *Avsunviroidae* (Fadda et al., 2003). The main features of this symptomless viroid are summarized in Table I.2. It has been proposed as the friendliest experimental system in the family *Avsunviroidae* (Daròs, 2016). This categorization is justified because eggplant is an herbaceous crop of economic importance, while, except chrysanthemum, all the other known hosts of this family are woody plants. Additionally, eggplant can be infected by viroids of the family *Pospiviroidae* such as PSTVd (Minoia et al., 2015; López-Carrasco et al., 2017) which makes it ideal for comparative analysis between the two families. Nonetheless, the eggplant genome has been recently re-sequenced (Hirakawa et al., 2014; Barchi et al., 2019; Gramazio et al., 2019) evidencing that it is a well-studied crop with ongoing attention. Thus, significant experimental work has been conducted with this viroid to gain insights into RNA replication (Nohales et al., 2012b), processing (Martínez et al., 2009; Cordero et al., 2018) and trafficking (Gómez and Pallás, 2010b; Gómez and Pallás, 2010c; Gómez and Pallás, 2012a). Some biotechnological applications have even been investigated (Daròs et al., 2018).

Table I.2 Main characteristics of ELVd

TAXONOMY:	*Eggplant latent viroid* is the only species of the genus *Elaviroid* (family *Avsunviroidae*).
PHYSICAL PROPERTIES:	ELVd consists of a single-stranded, circular RNA of 332–335 nucleotides that folds in a branched quasi-rod-like minimum free-energy conformation. Replication occurs in the plastids of infected cells through a symmetrical rolling circle.
HOSTS and SYMPTOMS:	ELVd has only been shown to infect eggplant (*Solanum melongena* L.) without causing any visible symptom. Initial analyses showed that ELVd is uniformly distributed in the leaves, stems and fruits (skin and pulp).
TRANSMISSION	ELVd is transmitted mechanically and by seed (approximately 20%).

The secondary structure of ELVd has been elucidated, resulting in similar bifurcated conformation for both polarity strands (Figure I.6). These structures have been supported by *in silico*, *in vivo* and *in vitro* data and the transcription initiation sites have been mapped in their *in vivo* conformation (Giguère et al., 2014a; López-Carrasco et al., 2016). Curiously, in the hammerhead ribozyme of the (+) polarity strand, the cleavage site is preceded by the trinucleotide AUA, which is absent in other natural hammerhead structures. The mutagenesis of this trinucleotide (to GUA, GUC and

AUC) resulted in *in vitro* self-cleavage rate constants higher than the wild-type at very low Mg^{2+} concentrations as found *in vivo* (Carbonell et al., 2006). Conversely, self-cleavage of the wild-type ribozyme (AUA trinucleotide) during transcription was more efficient than the mutants (GUA, GUC and AUC). These results led Carbonell *et al* to propose that viroid hammerheads have been evolutionarily selected to function co-transcriptionally to cleave the replication intermediates whilst protecting the viroid progeny (Carbonell et al., 2006).

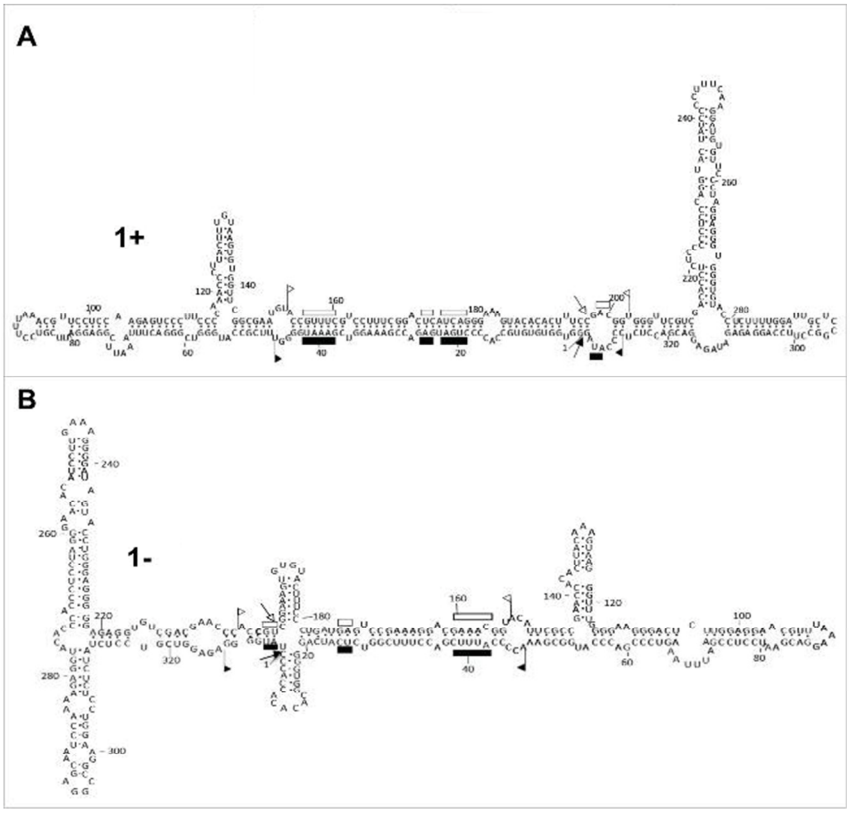


Figure I.6. Conformations of minimum free energy predicted for the (+) (A) and (-) (B) strands of ELVd. Sequences forming the hammerhead structures are delimited by flags, motifs conserved in natural hammerhead structures are denoted by bars, and self-cleavage sites are marked by arrows. Solid and open symbols refer to (+) and (-) polarities, respectively. The same numbers are used for both polarities. Figure reproduced from López-Carrasco *et al.* (2016).

Interestingly, the mutation rate of ELVd has been determined by high-fidelity ultra-deep sequencing using the lethal mutation method and the values per site ranged from 1/100 to 1/800 (López-Carrasco et al., 2017), relatively similar to 1/400 which was previously obtained for chrysanthemum chlorotic mottle viroid (CChMVd) (Gago et al., 2009). In contrast, those of PSTVd in the same host ranged from 1/7000 to 1/3800 depending on the sequencing run. These results reinforced the extremely high mutation rate as a common feature of chloroplastic viroids, whereas the mutation rates of PSTVd and potentially other nuclear viroids are significantly lower and closer to those of RNA viruses.

The molecular characteristics of ELVd (Table I.2) implied that it would replicate and accumulate in the plastids as it is assumed for the members of the family *Avsunviroidae*. Chloroplastic accumulation has been confirmed by *in situ* hybridization with digoxigenin-labeled RNA probes (Daròs, 2016). On the other hand, transplastomic lines of the green algae *Chlamydomonas reinhardtii* expressing dimeric viroid transcripts in the chloroplast were obtained to study mechanistic aspects of RNA processing (Molina-Serrano et al., 2007). Using this system, ELVd was more efficiently processed into mature circular forms than other family members, although viroid replication did not take place (Molina-Serrano et al., 2007). Moreover, the same transplastomic approach was used with ELVd mutants to determine that the hammerhead ribozyme-containing sequence was the only domain involved in cleavage, but additional sequences were involved in ligation (Martínez et al., 2009).

How avsunviroids are trafficked into the host chloroplasts has been an intriguing question (Daròs et al., 2006). In order to shed light on this aspect, Gómez and Pallás inserted a chimeric ELVd sequence that consisted of a fragment of the (−) strand (position 54–267, note that the numbering of the (-) strand goes backward), followed by a fragment of the (+) strand (position 54–261, including two mutations), immediately upstream of the green fluorescent protein (GFP) (Gómez and Pallás, 2010b). This chimeric ELVd sequence did not contain either complete viroid ribozymes that may mediate mRNA cleavage or an alternative AUG in frame with that of the GFP, and thus it was assumed that the fluorescence observed in the chloroplast was a result of the translation of the chimeric fusion in the organelle. Later, this experimental strategy was also extended to CChMVd, another member of the family, with similar results (Baek et al., 2017). However, a more recent approach trying to optimize this ELVd system for carrying RNAs to the chloroplast for biotechnological purposes has unexpectedly revealed that instead of being translated in the chloroplast, the GFP is transported to the organelle by a chloroplast transit peptide (cTP) derived from the viroid sequence (Márquez-Molins, J. 2017, Final Master Thesis). This viroid-derived peptide probably started in a non-canonical codon, a phenomenon already described in plant cells (Gordon et al., 1992) that could even be of biological significance (Simpson et al., 2010). Moreover, this cTP is presumably cleaved by the interaction with the TOC-TIC (translocon at the outer chloroplast envelope–translocon at the inner chloroplast envelope) import machinery as shown for canonical cTP of proteins trafficked to this organelle (Inaba and Schnell, 2008). This cleavage explains the reason why Gómez and Pallás did not observe differences in size between the GFP and the chimeric construct fused to GFP in a western blot analysis using 10% of acrylamide and therefore assumed that the chimeric sequence was acting as a non-coding RNA. A western blot analysis using gels of higher acrylamide percentage revealed that the band corresponding to the construct carrying the chimeric viroid actually showed bigger size than GFP

(27 kDa). Moreover, a mutant with two extra nucleotides between the chimeric ELVd and a yellow fluorescent protein (YFP) was analyzed. Confocal microscopy observations of this putative frameshift mutant revealed a nucleo-cytoplasmic distribution of the fluorescence, which was not observed at all in the chloroplasts.

A study regarding the intracellular trafficking of ELVd (Gómez and Pallás, 2012a) has been used to propose the interesting notion that a nuclear step may exist also in the family *Avsunviroidae* (Gómez and Pallás, 2013). The experimental system in which the ELVd sequence was incorporated, was based on potato virus X (PVX). This system had been previously used to demonstrate the nuclear targeting of PSTVd (Zhao et al., 2001), and consisted of a recombinant version of PVX carrying a GFP gene that contains the intron IV2 (derived from the potato gene ST-LS1). Thus, in that system, fluorescence can only be detected if the intron-containing GFP mRNA traffics into the nucleus for appropriate splicing. A construct with the (+) full-length monomeric ELVd sequence (AJ536613) embedded within the IV2 intron was assayed in *N. benthamiana* plants. Observation of systemically infected leaves revealed GFP fluorescence, suggesting that the ELVd RNA directs the cytoplasmic GFP transcript to the nucleus, where the intron can be efficiently removed. Furthermore, the ELVd sequence was dissected into three arbitrary regions (left, from position 15 to 181; upper, from position 89 to 294; and right, from position 187 to 333), and only the left region was sufficient for mediating the nuclear traffic. According to these findings, in a natural infection, ELVd would first move to the nucleus and later be trafficked to the chloroplast where it replicates and accumulates (Figure I.7). However, the function of this putative nuclear traffic is still unknown.

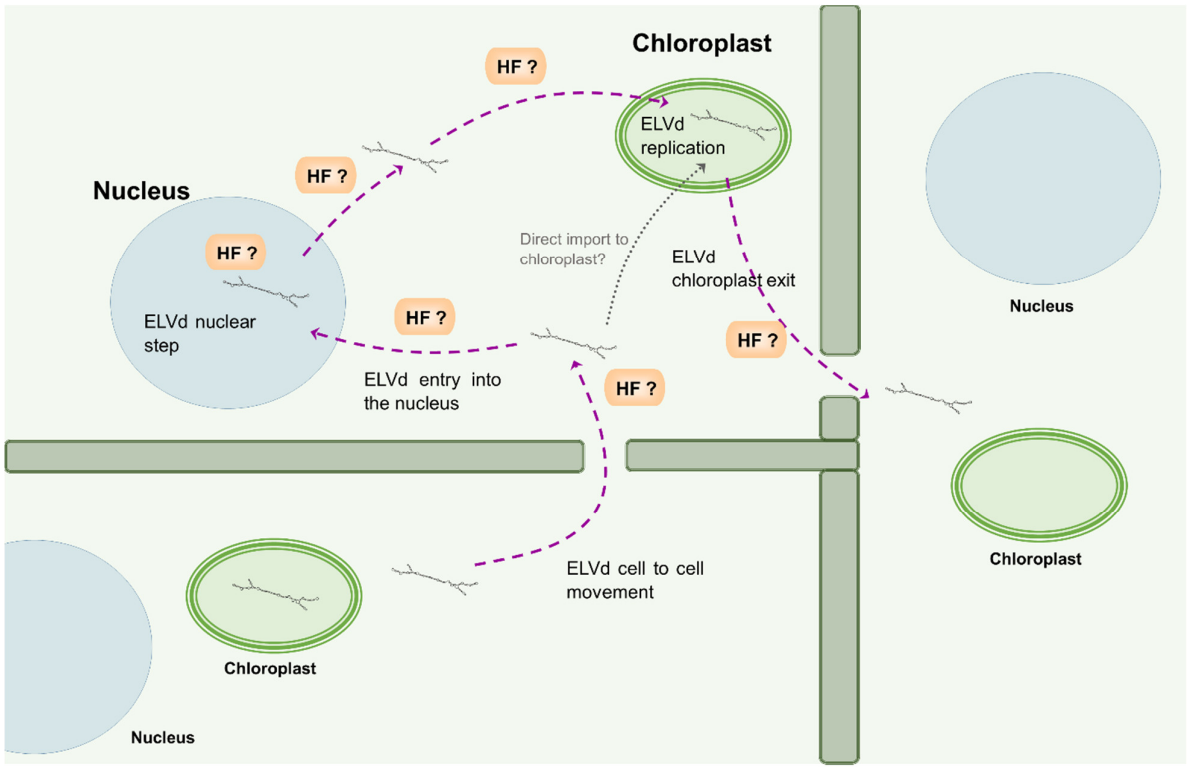


Figure I.7. Proposed model to explain the specific compartmentalization of ELVd into chloroplasts. In this transport pathway (likely mediated by an RNA domain localized in the left terminal region), after the viroid invades the cell cytoplasm, it is imported into the nucleus, by means of an unknown host-dependent mechanism. The viroid then uses this organelle as a sub-cellular shuttle for delivery into the chloroplast, where replication takes place. The possibility that avsunviroids can directly traffic from the cytoplasm to the chloroplast cannot be excluded. Figure adapted from Gómez and Pallás (2012b).

4. Host alterations induced by viroids

As naked RNA pathogens, viroids extensively rely on host factors for their movement and replication (Ding, 2009; Adkar-Purushothama and Perreault, 2020). Importantly, the infection process triggers an array of host responses that result in the phenotypic alterations identified as symptoms (Navarro et al., 2021a). At a transcriptomic level, these alterations have been characterized in several host-viroid pathosystems (Owens et al., 2012; Xia et al., 2017; Zheng et al., 2017; Thibaut and Claude, 2018; Góra-Sochacka et al., 2019; Štajner et al., 2019; Takino et al., 2019; Wang et al., 2019b). In general, genes with differential expression are involved in plant defense, hormone signalling, photosynthesis, RNA processing and binding, and metabolism (reviewed in Owens et al., 2017). It is worth noting that despite being naked RNAs, viroids can activate pathogen-associated molecular pattern (PAMP)–triggered immunity (PTI) and effector-triggered immunity (ETI) (Owens et al., 2017). In particular, nuclear-replicating viroids can induce the upregulation of genes involved in plant innate immune responses—such as pathogenesis-related proteins, mitogen-activated protein kinase, resistance proteins, defense-related transcription factors, and genes involved in cell wall biogenesis and reactive oxygen species metabolism (Zheng et al., 2017; Thibaut and Claude, 2018; Štajner et al., 2019; Takino et al., 2019). Additionally, it is particularly interesting how viroids directly interfere with both the RNA silencing machinery (Itaya et al., 2001; Papaefthimiou et al., 2001) and the epigenetic mechanisms that maintain the stability of the host genome (Martinez et al., 2014; Castellano et al., 2016b; Castellano et al., 2016a).

4.1 RNA silencing and viroids

RNA silencing is a sequence-specific mechanism that regulates gene expression in many eukaryotic organisms, including plants (Baulcombe, 2004). Diverse pathways producing transcriptional or post-transcriptional gene silencing have been described (Brodersen and Voinnet, 2006; Martínez de Alba et al., 2013). Because of their sessile nature, RNA silencing in plants is not only crucial for regulating developmental processes but also to cope with stress conditions, especially as a defense against viral and sub-viral pathogens (Ding and Voinnet, 2007; Islam et al., 2018). RNA silencing is an evolutionary conserved mechanism that is characterized by: i) formation of helical RNA precursors, such as double-stranded RNAs (dsRNAs); ii) processing of precursors to 20-24-nucleotide small non-coding RNA (sRNA); and iii) inhibitory action of sRNAs on complementary RNA or DNA (Brodersen and Voinnet, 2006). According to their biogenesis, endogenous plant sRNAs are classified into those derived from single-stranded RNAs (ssRNAs) precursors with a hairpin structure: hairpin RNAs (hpRNAs); or those derived from dsRNA precursors: small interfering RNAs (siRNAs) (Axtell, 2013). The production of both types of

sRNAs is the result of the activity of RNase III-type enzymes known as Dicer-like (DCL) and depending on the resultant 5' terminal nucleotide are loaded onto specific AGO that form the RNA induced silencing complex (RISC) (Vaucheret, 2008), which is responsible of the post transcriptional gene silencing (PTGS).

Viroids, as well as RNA viruses, trigger RNA silencing, which is evidenced by the presence of vd-sRNA in infected tissues with viroids of both families (Itaya et al., 2001; Papaefthimiou et al., 2001; Martinez de Alba et al., 2002; Hammann and Steger, 2012). In fact, it has been proposed a close interplay between viroid-induced pathogenesis and RNA silencing (Wang et al., 2004; Gómez et al., 2009). According to this, vd-sRNA act as elicitors of symptoms expression by cleaving endogenous mRNA of the host. Although initially controversial (Di Serio et al., 2009; Di Serio et al., 2010), the involvement of vd-sRNA as inductor of certain host symptoms is now unanimously accepted, as phenotypically-relevant targeted host transcripts have been identified in different plants infected with members of both families *Pospiviroidae* (Eamens et al., 2014; Adkar-Purushothama et al., 2015; Adkar-Purushothama et al., 2018; Bao et al., 2019) and *Avsunviroidae* (Navarro et al., 2012b; Delgado et al., 2019). The direct requirement of an active state of RNA silencing pathways for the apparition of symptoms was also experimentally validated in transgenic HSVd-Nb plants (Gómez et al., 2008). At low temperature (14 °C) RNA silencing is inhibited (Szittya et al., 2003; Qu et al., 2005) and HSVd-Nb plants were asymptomatic and expressed normally an HSVd-GFP reporter, whereas plants grown at high temperature (28 °C) developed clear symptoms of HSVd infection and the expression of the reporter was suppressed. Moreover, tomato transgenic plants expressing a hairpin-RNA transgene consisting of an incomplete but nearly full-length PSTVd sequence did not accumulate mature PSTVd RNAs but had phenotypes that closely resemble those of PSTVd infected plants (Wang et al., 2004). Furthermore, artificial microRNAs of viroid sequences resulted in phenotypes resembling, at least partially, those of viroid disease (Eamens et al., 2014; Adkar-Purushothama et al., 2018; Bao et al., 2019). Finally, it was reported that dsRNAs interfere in a sequence-specific manner with the infection of pospiviroids (PSTVd and citrus exocortis viroid, CEVd) and avsunviroids (CChMVd) (Carbonell et al., 2008). Therefore, viroid RNAs can be both substrates and targets of the RNA silencing machinery.

Whether vd-sRNAs are primarily originated from single or double-stranded RNAs has been a controversial issue. It was suggested that the unencapsidated and highly structured genomic viroid ssRNAs could be vulnerable to DCL enzymes (Landry and Perreault, 2005; Carbonell et al., 2008). However, it was independently reported, in PSTVd and HSVd, that mature viroids display a certain degree of resistance to RISC-mediated cleavage (Itaya et al., 2007; Gómez et al., 2008), presumably because of the high degree of secondary structure of these circular RNAs (Elena et al., 2009).

Moreover, deep sequencing analysis of vd-sRNAs and a more profound mechanistic knowledge of the RNA silencing machinery suggest a significant role of dsRNAs in the generation of vd-sRNAs in pospiviroids. First, vd-sRNAs of both polarities are present at significant levels in infected tissues. Second, the preferential mapping of those vd-sRNAs at hotspots was used as an argument of their ssRNA origin, but DCL-processing of even perfect dsRNAs can also produce hotspot regions in sRNAs populations (Sasaki et al., 2014). Additionally, in transgenic tomato plants expressing a hairpin-PSTVd transgene, the derived sRNAs showed hotspot distribution patterns similar to those that were detected in natural infections (Schwind et al., 2009). Furthermore, the parallelisms between the biogenesis of trans-acting small interfering RNAs (ta-siRNAs) and the life cycle of pospiviroids suggested the incorporation of viroid replication intermediates into the ta-siRNAs biogenesis pathway (Gómez et al., 2009).

Conversely, the biogenesis of vd-sRNAs in avsunviroids should be different as these viroids exclusively replicate in the chloroplast and therefore the silencing machinery theoretically could not have access to the replication intermediates (Navarro et al., 2000). Thus, vd-sRNA must be generated as a consequence of viroid traffic through the cytoplasm or the nucleus (Gómez and Pallás, 2013). It can be hypothesized that cytoplasmic RNA-dependent RNA polymerases, like RDR6, might be involved in the generation of dsRNAs from mature viroids. These dsRNAs could be substrates of DCL4 and DCL2 and that would explain the reported accumulation of 21- and 22-nt vd-sRNA, respectively (Dunoyer et al., 2005; Bouché et al., 2006). Importantly, the proportion of 24-nt vd-sRNAs of chloroplast-replicating viroids is extremely low, which suggests the lack of involvement of DCL3 (Navarro et al., 2012b).

The interactions with the silencing machinery have been studied more in detail in PSTVd using *N. benthamiana* knockdown lines of RDR6 (Di Serio et al., 2010) and DCL proteins (Dadami et al., 2013; Katsarou et al., 2016a), in addition to *in vivo* immunoprecipitation with epitope-tagged versions of AGO proteins (Minoia et al., 2014) and cytoplasmic and nuclear PTGS sensor constructs (Dalakouras et al., 2015). First, it was observed that RDR6 is associated with a delay of PSTVd accumulation and with its exclusion from the meristem (Di Serio et al., 2010). Later, the knockdown of DCL1 and DCL4 diminished viroid accumulation in systemic infection (3 weeks after inoculation, wpi), which was unaltered in DCL2 and DCL3 knockdowns (Dadami et al., 2013). However, in protoplasts transfected with dimeric PSTVd RNA, the suppression of DCL2 and DCL4 increased viroid levels, while DCL1 or DCL3 suppression does not seem to significantly affect viroid titer. Therefore, although DCL4 processing of PSTVd may decrease its titer on a single plant cell, it would eventually facilitate viroid's systemic spread, which reflects a remarkable evolutionary adaptation of viroids to use the silencing machinery in their benefit (Dadami et al.,

2013). Moreover, simultaneous knockdown of DCL1, DCL2 or DCL3 together with DCL4 could not reverse the titer decrease (Katsarou et al., 2016a). Thus, it has been proposed that viroids, like ta-siRNAs (Howell et al., 2007), might have evolved to be primarily processed by DCL4 (Gómez et al., 2009; Katsarou et al., 2016a). Conversely, the combined suppression of DCL2 and DCL3 had a strong positive effect on PSTVd accumulation, indicating the likely role of both proteins in anti-viroid defense (Katsarou et al., 2016a). On the other hand, all AGOs except AGO6, AGO7, and AGO10 were associated with PSTVd-sRNAs: AGO1, AGO2, and AGO3 preferentially with those of 21 and 22 nt, while AGO4, AGO5, and AGO9 additionally bound those of 24 nt (Minoia et al., 2014). Deep-sequencing analyses showed that sorting of vd-sRNAs into AGO1, AGO2, AGO4, and AGO5 depended essentially on their 5′-terminal nucleotides (Minoia et al., 2014). Furthermore, the overexpression of AGO1, AGO2, AGO4, and AGO5 attenuated viroid accumulation, supporting their involvement in plant defense against viroids (Minoia et al., 2014). Finally, PTGS predominantly occurs in the cytoplasm, but nuclear PTGS has been also reported (Hoffer et al., 2011). However, using specific sensors it was shown that PSTVd only triggers cytoplasmic PTGS, despite the co-localization in the nucleus of plant DCLs and pospiviroid RNA intermediates (Dalakouras et al., 2015). This result was in agreement with the previous observation of PSTVd-siRNAs in the cytoplasm (Denti et al., 2004).

Taken together, these results indicate a complex equilibrium by which viroids have overcome to a certain extent plant RNA silencing defensive mechanisms, adopting highly structured forms and specific compartmentalization, but are still susceptible in certain conditions, which is evidenced by the presence of vd-sRNAs in infected tissues. In the case of nuclear-replicating viroids, the biogenesis of vd-sRNAs probably follows primarily the RDR6-DCL4 pathway, while the situation for *Avsunviroidae* members is much less clear, as the inability to infect model plants has hindered research in this area. What seems even more interesting is that viroids could benefit from the interaction with the RNA silencing machinery. For example, in plants infected with PSTVd, callose synthase genes are targeted, which may facilitate viroid cell to cell movement through plasmodesmata (Adkar-Purushothama et al., 2015) and early flowering is also induced by PSTVd-sRNAs, which could be beneficial for PSTVd dispersal through seeds (Adkar-Purushothama et al., 2018). Moreover, the only two host transcripts currently identified to be targeted by vd-sRNAs of a chloroplast-replicating viroids code for chloroplastic proteins (Navarro et al., 2012b; Delgado et al., 2019).

4.2 Epigenetic alterations caused by viroids

Plant responses to biotic stress involve a series of coordinated signalling networks to identify and respond to the pathogen invasion by minimizing the damage at the expense of development (Pumplin and Voinnet, 2013). The switch from the regular developmental program to the stress response program requires changes in gene expression that produce proteins to cope with stress and inhibit growth-related pathways. This is termed transcriptional reprogramming and is the consequence of dynamic changes at a deeper level in the genome (Tsuda and Somssich, 2015). Plants have widespread epigenetic marks that are responsible for controlling gene expression (Slotkin, 2016). These modifications can be classified as microstructurals (at the DNA level) and macrostructurals (at the chromatin level) (Köhler and Springer, 2017; Annacondia et al., 2018). DNA modifications in plant genomes mostly consist of methylation of the fifth carbon of cytosine residues (5mC) and can occur in all three different contexts (CG, CHG, CHH; where H = A, C or T), (Zhang et al., 2018a). Overall, the DNA methylation state is the outcome of dynamic regulation *by de novo* methylation, maintenance mechanisms and active demethylation, processes which are catalysed by various enzymes and targeted by distinct regulatory pathways (Zhang et al., 2018a). Specifically, cytosine methylation is established *de novo* by siRNAs through the pathway denominated RNA-directed DNA methylation (RdDM), and generally results (but not always) in compacted chromatin and transcriptional gene silencing (TGS) (Matzke et al., 2015). RdDM is guided by small RNAs predominantly of 24 nt in size, derived from double-stranded precursors synthesized by RNA Polymerase IV (Pol IV), in close partnership with the RNA-dependent RNA polymerase 2 (RDR2) (Wendte and Pikaard, 2017). On the other hand, the chromatin structure is additionally influenced by histone marks such as acetylation, methylation, and ubiquitination. Interestingly, these histone modifications can have different effects in the transcriptional status despite their chemical similarity. For instance, in *Arabidopsis*, histone H3 lysine 4 (H3K4) mono−/di- or tri-methylation is associated with highly transcribed genes (Zhang et al., 2009) while H3K27 tri-methylation is found in silenced genes (Zhang et al., 2007), and H3K9 di-methylation is predominantly present in silenced transposable elements (TE) (Zhou et al., 2010). In short, host transcriptional reprogramming is a central part of plant defense upon pathogen recognition and it is tightly regulated by epigenetic mechanisms that modify the accessibility of the transcription machinery (Tsuda and Somssich, 2015; Annacondia et al., 2018).

Increasing evidence links the pathogen invasion with epigenetic modifications in the host genome (Gómez-Díaz et al., 2012; Silmon De Monerri and Kim, 2014). DNA viruses are known to interfere with the host DNA methylation machinery (Rodríguez-Negrete et al., 2013; Yang et al., 2013) to avoid the methylation of their genome, and the consequent transcriptional repression (Raja et al.,

2008). This alteration is mediated by proteins that suppress silencing and thus is related to the transcriptional reactivation of TE in the host genome (Rodríguez-Negrete et al., 2013; Yang et al., 2013). Similarly, RNA viruses generally code for silencing suppressors (Csorba et al., 2015), and, for example, it has been demonstrated that those of cucumber mosaic virus can interact with siRNAs from the RdDM pathway altering the methylation levels of its target genes (Hamera et al., 2012).

Even though viroids do not code for any protein capable of suppressing silencing, the existence of a relationship between viroid infection and DNA methylation has been proposed for a long time (Wassenegger et al., 1994; Dalakouras et al., 2013). In fact, it led to the discovery that a mechanism of *de novo* DNA methylation directed by RNA existed in plants since the full methylation of the cDNA of PSTVd in transgenic plants was only observed after autonomous viroid RNA-RNA replication had taken place (Wassenegger et al., 1994). This was the first evidence that RdDM existed in any organism and was crucial to the research on this topic (see review by Wassenegger and Dalakouras, 2021).

Further studies with PSTVd revealed that infection with this viroid in potato caused the hypermethylation of promoter regions of certain genes, and the consequent transcriptional repression (Lv et al., 2016). Moreover, transient expression of PSTVd in GFP transgenic in *N. benthamiana* plants (line 16c) accelerated GFP silencing by increasing DNA methylation in the promoter of this transgene (Lv et al., 2016). Interestingly, PSTVd increased the expression of a protein bromodomain-containing viroid-binding protein (Virp1) which is required for PSTVd replication (Kalantidis et al., 2007), and the transient overexpression of Virp1 also increased methylation levels of the GFP transgene, but to a lesser extent in comparison with PSTVd. Additionally, it was revealed that PSTVd produces a strong activation of host DNA methylation pathways in tomato (Torchetti et al., 2016). In particular, PSTVd upregulated the expression of key genes involved in the maintenance of DNA methylation, such as methyltransferase 1 (MET1), chromomethylase 3 (CMT3) and domains rearranged methyltransferase 2 (DRM2) and also related with histone methylation like H3K9 histone methyltransferase KRYPTONITE/SUVH4 (KYP). The strong enhancement of the DNA methylation machinery triggered by PSTVd resulted in clear antagonism with an ssDNA geminivirus, which was hypermethylated in mixed infections of these two nuclear-replicating pathogens (Torchetti et al., 2016). All these results support that broad DNA methylation may be enhanced by PSTVd infection. However, how PSTVd could induce overexpression of methylation-related genes is still a mystery.

Conversely, HSVd infection induces the hypo-methylation of the promoters of rRNA genes presumably by the subversion of HDA6 (Martinez et al., 2014; Castellano et al., 2015; Castellano

et al., 2016b). HDA6 is responsible for the deacetylation of lysine residues on the N-terminal part of the core histones (H2A, H2B, H3 and H4) and gives a tag for epigenetic repression (Liu et al., 2012b; Liu et al., 2012a). It has been associated with the removal of acetyl residues from specific targets, such as rDNA repeats or complex transgenes (Liu et al., 2012a). Therefore, the recruitment by HSVd functionally inactivates host HDA6, and that results in the transcription of ribosomal genes. Overexpression and suppression of HDA6 gave evidence for a functional link between the activity of this enzyme and HSVd biological efficiency (Castellano et al., 2016b). Interestingly, the lack of HDA6 activity has been associated with spurious Pol II transcription of nonconventional rDNA templates (usually transcribed by RNA Pol I) (Earley et al., 2010). In HSVd-infected plants, it was reported an overaccumulation of preribosomal RNAs (pre-rRNAs) and sRNAs derived of ribosomal RNA (rb-sRNAs) indicating an unusual transcriptional environment (Martinez et al., 2014; Castellano et al., 2015; Castellano et al., 2016b). Therefore, it was proposed that the HDA6-recruitment mediated by HSVd may promote spurious RNA Pol II activity which could favour the transcription of non-canonical templates as viroid genomic RNA.

In conclusion, nuclear-replicating viroids can significantly alter the epigenetic regulation of the host genome, but up to date, only particular effects have been reported and the alterations on a genomic scale are unknown. It could be hypothesized that different viroids of the family *Pospiviroidae* might cause radically opposite effects by interacting with different host proteins, based on the hypermethylation and hypomethylation phenomena described in PSTVd and HSVd, respectively. However, both effects may not be incompatible, as viroids could inactivate HDA6 and increase the transcription of other methylation-related genes to cause a generalized inhibition of transcription and thus facilitate the recognition of non-canonical templates by pol II, not only by the inactivation of HDA6 but also because of a decrease of the genomic regions accessible to be actively transcribed. Moreover, although no epigenetic alteration has been described to be caused by avsunviroids, it cannot be discarded since a nuclear step has been proposed for this family (Gómez and Pallás, 2013).

5. The contribution of viroids to molecular biology

The singularities of viroids, such as being the smallest infectious nucleic acids known so far, having highly structured circular RNA genomes, the existence of hammerhead ribozymes embedded in *Avsunviroidae* sequences or their apparent non-coding capacity, are behind some relevant discoveries in the field of molecular biology (Figure I.8). The study of viroids was pioneering and initiated the development of innovative methods, such as chromatographic or electrophoretic, which were later used in molecular biology, gene technology and in prion research (Steger and Riesner, 2018). Moreover, viroids have been proposed as useful experimental models for studying many aspects of RNA biology (Diener, 1989; Ding and Itaya, 2007; Ding and Wang, 2009).

In 1971, Diener described the first viroid as a replicating, low molecular weight RNA responsible for the spindle tuber pathology (Diener, 1971). A few years later, the circularity of viroid RNA was observed by electron microscopy, constituting the first detection of a circular RNA in nature (Sanger et al., 1976). This circularity was confirmed by the sequencing of PSTVd, a significant milestone, as no other pathogen infecting a eukaryotic host had been sequenced before (Gross et al., 1978). Theoretical and experimental studies of viroid RNA folding demonstrated the biological importance of metastable structures (Henco et al., 1977; Henco et al., 1979; Riesner et al., 1979). This pioneering knowledge was later applied to other RNAs and was used to establish the principles of structure formation emerging from folding kinetics (Steger and Riesner, 2018). Moreover, the analysis of a viroid fragment by nuclear magnetic resonance (NMR) spectroscopy demonstrated directly that hydrogen bonds in RNA have partially covalent character (Dingley and Grzesiek, 1998; Dingley et al., 2003). This evidence has had implications in the establishment of unambiguous hydrogen-bond networks in secondary and tertiary structures of DNA and RNA, as well as in the study of H-bond cooperativity (Steger and Riesner, 2018).

Even though self-catalytic propriety had been previously described in introns (Cech et al., 1981), the discovery of hammerhead ribozymes in the genome of a pathogen (Hutchins et al., 1986), and viroid secondary structures in general, were important to reinforce the relationship between RNA structure and biological function. More recent studies have found hammerhead ribozymes in genomes all across the tree of life and transcriptomic evidence of these self-catalytic motifs in highly expressed circular RNAs (circRNAs) (de la Peña et al., 2017; Cervera and De La Peña, 2020). Interestingly, these circRNAs have been suggested as remnants of a prebiotic RNA world (Diener, 1989; Flores et al., 2014). Besides the putative involvement in the early stages of life, there is a growing interest in the functional role that circRNAs could be playing in nowadays organisms (Patop et al., 2019).

CircRNAs are characterized by a prolonged half-life in comparison with their linear counterparts and have been attributed functions related to the regulation of gene expression (Memczak et al., 2013; Patop et al., 2019). Despite initially being considered as non-coding RNAs (Guo et al., 2014), recent evidence supports the translation of certain circRNAs in eukaryotic cells (Legnini et al., 2017; Pamudurti et al., 2017) and biotechnological applications are even being investigated (Wesselhoeft et al., 2018). Therefore, the true coding potential of circRNAs may have been underestimated, which could be a justification to re-examine the categorization of viroids as non-coding RNAs. Viroids share some characteristics with endogenous circRNAs like the dependence of RNA pol II, the size range (150-600 bp) and non-canonical RNA processing. Despite many circRNAs have been identified in plants, the knowledge about their functional roles is very scarce and, for instance, have a lower prediction to be potential miRNA sponges than those of animals (Ye et al., 2015). Thus, exploring the coding potential of viroids could have implications in expanding the plant transcriptome to circRNAs.

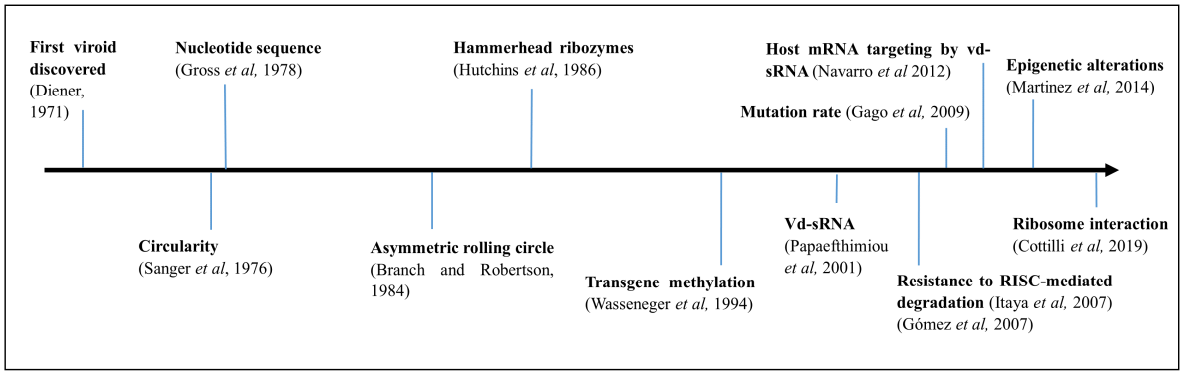


Figure I.8. Milestones in the molecular biology of viroids.

Viroids contain in their extremely small genome all the information necessary to mediate replication, intra- and intercellular trafficking, systemic movement and pathogenicity. This structural simplicity, functional versatility, and experimental tractability of viroid RNAs turn them into excellent models to investigate basic principles of infection (Ding and Itaya, 2007). For example, two viroids co-infecting the same host were used for the experimental confirmation of the quasi-species effect, by which, at high mutation rates, the fittest organisms are not necessarily the fastest replicators but rather are those that show the greatest robustness against deleterious mutations (Codoñer et al., 2006). Moreover, it was predicted that there is an inverse correlation between error rate and genome size, which was in part confirmed by the determination of the error rates of a viroid (Gago et al., 2009).

RNA movement is a phenomenon still poorly understood and the specific compartmentalization of these pathogens is valuable for studying intracellular, intercellular, and systemic RNA trafficking, as it has been exemplified with HSVd, ELVd and PSTVd (Ding and Wang, 2009; Gómez and Pallás, 2013; Pallas and Gómez, 2013). Studying the viroid cell cycle has given insights into RNA-template replication (Branch and Robertson, 1984) and how both families of viroids subvert DNA polymerases to accept RNA as a template (Mühlbach and Sänger, 1979; Warrilow and Symons, 1999; Navarro et al., 2000; Rodio et al., 2007). Additionally, viroids have been crucial for studying RNA-based mechanisms that regulate gene expression. The notion that RNA could direct DNA methylation was proposed to explain the methylation of a PSTVd transgene observed as a consequence of PSTVd infection and was named RNA-directed DNA methylation (Wassenegger et al., 1994). This phenomenon was reported at the beginning of the 90s, before any mechanism of RNA silencing was described (Fire et al., 1998). The discovery of RNA silencing has changed our understanding of how gene expression is regulated in eukaryotic genomes. Furthermore, RdDM is primordial for keeping genomic stability in plant genomes by inducing transcriptional silencing at repetitive DNA, including all types of transposons.

More recently, studies in tomato have shown that CEVd, which belongs to the same genus as PSTVd, and therefore to the family *Pospiviroidae*, causes ribosomal stress in host plants. In particular, CEVd is bound to the ribosome and causes alterations in both the processing of the 18S rRNA and in the translational profiles of tomato plants (Cottilli et al., 2019). Therefore, it is the first pathogen known to impair ribosome biogenesis.

Taken together, these findings evidence that the study of viroids has been relevant for molecular biology. The interest in these pathogens is not only fuelled by the economic losses which can cause in several crops, but also because of their role as experimental models to address fundamental questions in molecular biology. Importantly, the implementation of novel technologies in viroid research may make possible new contributions.

OBJECTIVES AND JUSTIFICATION

Since viroids were discovered in the seventies, we have gained considerable knowledge about their nature and replication mechanisms in host plants. However, many aspects of viroid biology are yet to be discovered. For example, how viroids cause symptoms in their hosts had been a conundrum, and only when silencing mechanisms have been discovered it has been possible to establish some direct relationships between viroid molecules and certain host symptoms, although we still have not a complete understanding of the process. Moreover, even though some factors involved in viroid movement have been discovered, we are far from deciphering the detailed mechanisms of RNA movement in plants that are hijacked by these intriguing pathogens. In general, this thesis aims to use innovative approaches for gaining a better understanding about viroids and the responses that they trigger in host plants. The particular objectives that are encompassed in that holistic vision and their justification are the following:

1. Development of an efficient method for constructing infectious clones of any viroid.

In order to produce artificial viroid infections, dimeric cDNA clones of the entire viroid sequence are used to generate transcripts that mimic the longer-than-unit replication intermediates and therefore can be processed into unitary, circular RNAs within the cell (Cress et al., 1983; Visvader et al., 1985). The first viroid cDNA infectious clone was developed in 1983 (Cress et al., 1983) and since then, constructs of this type have been used to study viroid pathogenesis in natural and experimental hosts. The use of these clones instead of nucleic acid extracts from infected plants has several advantages. First, the cDNA of viroid dimers provides a near-homogenous inoculum, whereas if isolated from infected plants the viroid RNA consists of a mix of sequence variants. Moreover, directed modifications in the viroid sequence can be performed, which enabled the identification of motifs or domains that are important for viroid replication and/or movement (Flores et al., 1999; Qi and Ding, 2002; Zhu et al., 2002; Qi et al., 2004), and also the characterization of a domain that interacts with a certain host factor (Gozmanova et al., 2003). In short, dimeric viroid cDNAs are essential tools for viroid research. Until the present work, dimeric clones of viroids were constructed by self-ligation of monomeric cDNA in presence of T4 DNA ligase (Cress et al., 1983; Adkar-Purushothama et al., 2015). However, given the inherent drawbacks of that approach (only a small fraction forms an exactly oriented dimer), the first objective of this thesis was to develop an oriented and therefore, more efficient and straightforward way for the construction of infectious viroids.

2. Exploring the coding potential of viroids of both families.

The recent discovery that circRNAs in the size range of viroids are translated in several eukaryotes (Diallo et al., 2019), made invalid the arguments of circularity and small size for considering viroids as non-coding RNAs. Moreover, some hypothesis may be proposed to explain why the experiments using cell-free systems in the seventies failed to reveal any protein derived of viroids. A first explanation would rely on the low sensitivity in the detection methodology. These experiments were carried out comparing with viral RNAs. Secondly, it cannot be excluded that viroid translation might be dependent on some host factors that were not present in the cell-free extract. Inspired by the recent findings of translating circRNAs in animals the second objective of this thesis was to re-examine the potential of viroids to serve as substrates for translation.

3. Characterization of host alterations caused by viroid infection at multiple levels.

The advent of the high-throughput sequencing technologies has favoured the study of the viroid-host interactions at a molecular level. However, such experimental approaches have been generally centred on the analysis of the viroid-induced changes at a single host-regulatory level and/or considering a particular infection time (mainly correlated to advanced disease stages). Therefore, our knowledge about the functional diversity and temporal dynamics of the global plant response to viroid infection is scarce (Adkar-Purushothama and Perreault, 2020; Marquez-Molins et al., 2021). In order to shed light on this aspect, the third chapter of this thesis consists of an integrative analysis of the timing and rate of genome-wide gene expression in response to HSVd infection. This viroid was the first known to cause an epigenetic alteration, the hypomethylation of the ribosomal gene promoter (Castellano et al., 2016b), but the genome-wide extent of the alterations triggered by HSVd was unknown. Thus, the objective of this analysis is to characterize the temporal dynamics of the plant response to viroid-infection at three different regulatory levels: i) small RNA interference, ii) modulation of the transcriptional activity and iii) epigenetic modifications.

CHAPTER 1

Highly efficient construction of infectious viroid-derived clones

Joan Marquez-Molins[1,2], José Antonio Navarro[2], Vicente Pallas[2] and Gustavo Gomez[1*].

[1] Institute for Integrative Systems Biology (I2SysBio), Consejo Superior de Investigaciones Científicas (CSIC) - Universitat de València (UV), Parc Científic, Cat. Agustín Escardino 9, 46980 Paterna, Spain.

[2] Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de Investigaciones Científicas (CSIC) - Universitat Politècnica de València, CPI 8E, Av. de los Naranjos s/n, 46022 Valencia, Spain.

[*] *Corresponding author*: E-mail: gustavo.gomez@csic.es

Published as Plant Methods 2019, 15, 87. doi:10.1186/s13007-019-0470-4.

PERSONAL CONTRIBUTION

I designed the cloning strategy, performed all the experiments, analyzed the results, and wrote the manuscript text.

ACKNOWLEDGEMENTS

We thank Alberto Carbonell for kindly providing pMDC32B-AtMIR390a-Bc vector. We thank Lorena Corachán for her excellent technical assistance. This work was supported by the Spanish Ministry of Economy and Competitiveness (co-supported by FEDER) Grants BIO2017-88321-R (VP) and AGL2016-79825-R (GG). I have been recipient of a predoctoral contract from the ACIF programme (ACIF-2017-114) of the Conselleria d'Educació, Investigació, Cultura i Esport Generalitat Valenciana.

Abstract and keywords

ABSTRACT

Viroid research generally relies on infectious cDNA clones that consist of dimers of the entire viroid sequence. At present, those dimers are generated by self-ligation of monomeric cDNA, a strategy that presents several disadvantages: i) low efficiency, ii) it is a non-oriented reaction requiring tedious screenings and iii) additional steps are required for cloning into a binary vector for agroinfiltration or for *in vitro* RNA production. We have developed a novel strategy for simultaneous construction of a viroid dimeric cDNA and cloning into a multipurpose binary vector ready for agroinfiltration or *in vitro* transcription. The assembly is based on IIs restriction enzymes and positive selection and supposes a universal procedure for obtaining infectious clones of a viroid independently of its sequence, with a high efficiency. Thus, infectious clones of one viroid of each family were obtained and its infectivity was analyzed by molecular hybridization. This is a zero-background strategy for direct cloning into a binary vector, optimized for the generation of infectious viroids. As a result, this methodology constitutes a powerful tool for viroid research and exemplifies the applicability of type IIs restriction enzymes and the lethal gene *ccd*B to design efficient and affordable direct cloning approaches of PCR products into binary vectors.

KEYWORDS

Viroid, cloning, dimers, IIs enzymes, agro-infiltration.

Introduction

Viroids are small single-stranded plant-pathogenic RNAs, being considered the smallest (246-434 nt) autonomous infectious nucleic acids known so far (Flores et al., 2015). These pathogens replicate and move systemically in host plants causing phenotypic effects that range from severe symptoms to latent infections (Flores et al., 2005; Gómez et al., 2009). More than 50 species of viroids have been described, being currently grouped into the families *Pospiviroidae* and *Avsunviroidae*, based on their replication site (nuclei and chloroplasts, respectively), presence of particular sequence domains, and properties of their infectious cycle (Di Serio et al., 2014). In both groups, replication takes place through a rolling-circle mechanism. Host RNA polymerases transcribe longer-than-unit replication intermediates that are cleaved and ligated to form circular monomers (Branch and Robertson, 1984). In nuclear viroids, only the (+) RNA intermediate is processed, whereas in members of the *Avsunviroidae* family both (+) and (−) intermediates are self-cleaved by hammerhead ribozymes and ligated subsequently (Daròs et al., 1994).

The particularities of viroids, such as being the smallest known pathogens, their extremely high mutation rate (Gago et al., 2009) or the fact that were the first circular RNAs discovered (Steger and Riesner, 2018), highlight the importance of the research on this topic. In this regard, novel methods have been developed by working on viroids, resulting in important applications such as determination of RNA folding and base pairing networks or setting the basis of plasmid purification kits (Steger and Riesner, 2018). Nonetheless, the study of viroids might still help to elucidate key biological pathways in plants such as RNA traffic (Gómez et al., 2004; Takeda et al., 2011; Gómez and Pallás, 2012a) or genetic regulation by epigenetic modifications (Wassenegger et al., 1994; Martinez et al., 2014; Castellano et al., 2016a).

The study of viroid replication and pathogenesis extensively relies on infectious cDNAs clones which consist of dimers of the entire viroid sequence whose transcripts mimic the longer-than-unit replication intermediates and therefore can be processed into unitary, circular RNAs within the cell (Cress et al., 1983; Tabler and Sänger, 1985). Due to this circular nature of viroidal genome, dimeric constructs are used in order to ensure that a full monomeric RNA can be produced independently of the cloning site. In this sense, experimental results indicate that viroid dimeric cDNAs always present a higher infectivity compared to monomeric cDNAs (Cress et al., 1983; Visvader et al., 1985). Thus, viroid dimeric cDNAs can be used for *in vitro* transcription and subsequent plant inoculation. However, the dimeric cDNA can be more easily and efficiently delivered to plant cells by agrobacterium-mediated transient plant expression, which avoids the requirement of obtaining a sufficient quantity of viroid transcript *in vitro*, and therefore has been

established as a more convenient strategy for inducing viroid-infection in diverse host plants (Gardner et al., 1986; Minoia et al., 2015; López-Carrasco et al., 2017).

The considerable decrease of the DNA synthesis price is supposing a revolution to the technology of recombinant DNA. However, the presence of repeated sequences is still an important limitation that often makes impossible the synthesis by commercial suppliers. Consequently, generating infectious viroids, which requires the dimerization of the sequence, and therefore intrinsically implies the generation of repetitions, is especially problematic.

At present, infectious dimeric viroids are constructed by self-ligation of monomeric cDNA in presence of T4 DNA ligase (Cress et al., 1983; Giguère et al., 2014b; Adkar-Purushothama et al., 2015). Dimeric head-to-tail clones on the desired (+) orientation are identified by restriction fragment analysis or using specific primers in colony PCR. However, the ratio of positive clones on the desired orientation could be scarce since self-ligation of blunt ends fragments is a non-controlled process in which only a small fraction forms an exact oriented dimer. Moreover, additional steps are required to introduce the dimer under an expression cassette in a binary vector for agroinfiltration or in a vector harboring T7/T3 polymerase promoters for *in vitro* transcription (Gardner et al., 1986).

In order to avoid tedious screenings and reduce timework, we have developed an oriented assembly strategy for efficient generation of viroid dimeric cDNAs. Assembly of DNA parts was introduced by Gibson Assembly and has been adapted to several methods (Gibson et al., 2009). Gibson assembly relies on generating cohesive ends and can produce a seamless fusion if overlapping DNA fragments are used. Those extremes can be generated by exonuclease activity being filled the gaps by a DNA polymerase, or alternatively using type IIs restriction enzymes, such as BsaI or BsmBI, which have non-palindromic recognition sites that are distal from the cleavage site (N1/N5) (Engler et al., 2009). Taking advantage of this property, we have constructed dimeric infectious constructs of one viroid of each family developing a novel strategy of simultaneous dimerization and cloning into a binary vector specifically designed for *Agrobacterium*-mediated viroid-inoculation.

Results

Dimeric clone construction

In order to enable the direct cloning of viroid dimeric cDNAs, a suitable vector for generating transcripts was specifically designed. This multipurpose vector includes a T7 RNA polymerase promoter for *in vitro* transcription and a double constitutive promoter 35S for agrobacterium-mediated transient plant transformation (Figure 1.1). The location of the T7 promoter, upstream the 35S, generates transcripts of the plus polarity of the viroid RNA (defined as the most abundant *in vivo*) which is the convenient to establish a viroid infection.

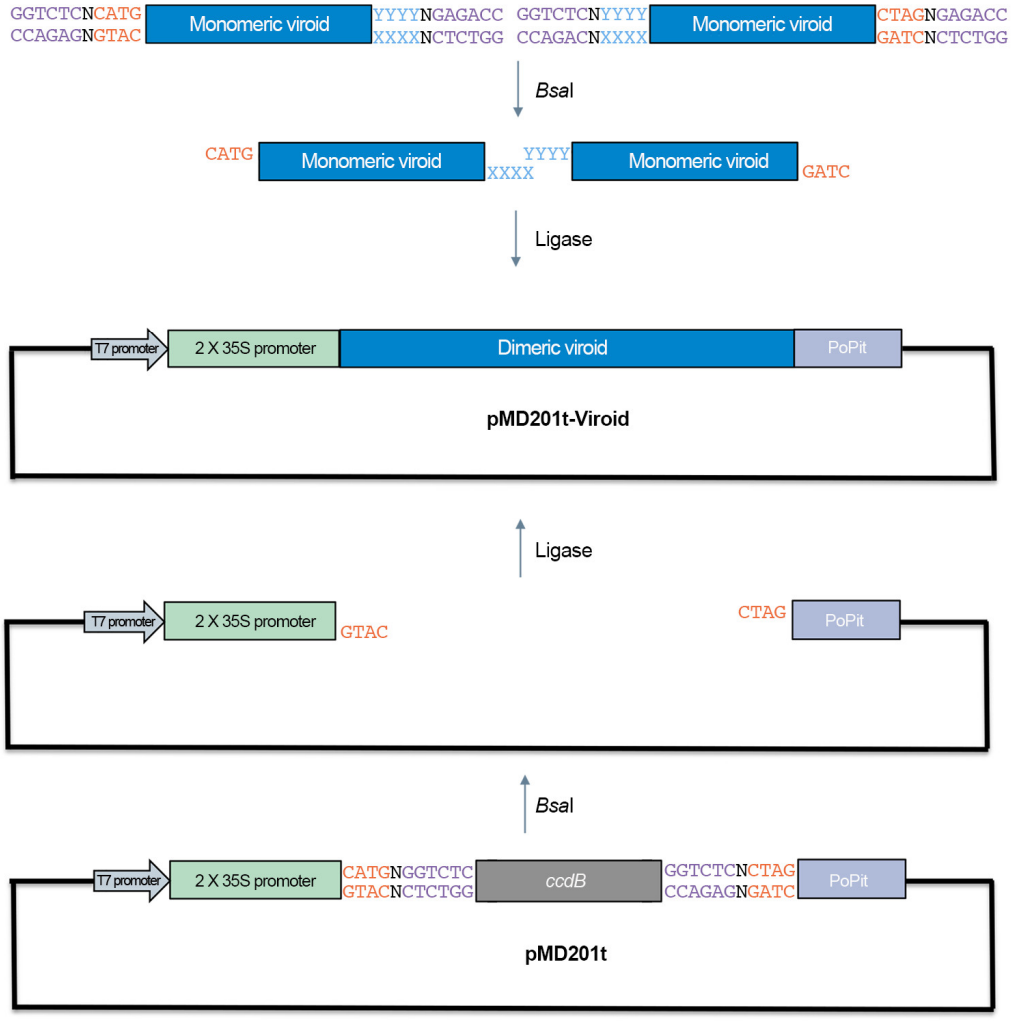


Figure 1.1. Schematic representation of viroid dimerization and subsequent assembly in a binary vector. A) Two different pairs of primers are used to generate distal BsaI recognition sites (purple) in such a way that compatible ends for assembly can be obtained. **B**) In a single reaction of simultaneous restriction and ligation the two viroid monomers (blue) are ligated between them and to a binary vector with compatible cohesive ends (orange). **C**) Specifically, the viroid dimer is inserted into an expression cassette that contains and a duplicated 35S, a constitutive promoter for plant expression and a PoPit terminator. Additionally, this binary vector has a T7 promoter for *in vitro* transcription. The dimeric viroid cDNA sequence replaces a lethal gene *ccd*B, thus guaranteeing an optimal efficiency of the reaction. **D**) Detail of the generation of the receptor-vector starting from pMD201t construct.

The resulting binary vector was named pMD201t because it was engineered from pMDC32B (Carbonell et al., 2014) and includes the same lethal gene (*ccd*B) for positive efficient selection. In pMD201t, *ccd*B is excised by BsaI digestion, generating four nucleotides overhang in each strand (GTAC in the negative strand and CTAG in the positive strand). The generation of these cohesive ends enables the orientated assembly of a desired cDNA designed to generate compatible ends upon digestion.

Experimental validation of the functionality of pMD201t was performed with a member of the *Pospiviroidae* family, hop stunt viroid (HSVd), in which BsaI is a non-cutter and with an *Avsunviroidae* member, eggplant latent viroid (ELVd), which contains a BsaI recognition site, thus exemplifying the applicability of the strategy for obtaining infectious clones of phylogenetically unrelated viroids.

The dimeric cDNA of HSVd was obtained as depicted in the Figure 1.1. Briefly, the viroid monomer was amplified using two pairs of primers (Fw D1-HSVd/Rv D2-HSVd and Fw D3-HSVd/Rv D4-HSVd, Supplementary Table 1.1) designed to generate, after BsaI digestion, cohesive ends that result in the merge of two monomers whose extremes are compatible with BsaI-digested pMD201t. Therefore, a seamless fusion of the HSVd dimer to pMD201t can be generated (the resultant construct was denominated pMD201t-HSVd). Taking advantage of the fact that IIs enzymes are active in ligase buffer (Engler et al., 2009), a one-pot reaction of restriction/ligation was set (Figure 1.2).

On the other hand, dimeric ELVd was generated following an equivalent strategy, except for the requirement of a previous step of digestion and purification prior to ligation to the binary vector pMD201t (Figure 1.2). Primers (Fw D1-ELVd/Rv D2-ELVd and Fw D3-ELVd/Rv D4-ELVd, Supplemental Table 1.1) were designed analogously for amplifying the viroid monomer but containing BsmBI recognition sites to produce the cohesive ends. After amplification and digestion with BsmBI, DNA fragments were ligated to pMD201t (the obtained construct was denominated pMD201t-ELVd).

The resulting ligations were transformed into *ccd*B-sensitive *E. coli* cells (DH5 Alpha), which lack the gene *ccd*A for producing the antitoxin. As a consequence, plasmid molecules in which *ccd*B has not been replaced cannot propagate, which results in a zero-background positive selection (Supplementary Figure 1.1). In this regard, colony PCR was performed to validate the presence of the viroid cDNA dimer in eight colonies for each construct (Supplementary Figure 1.2).

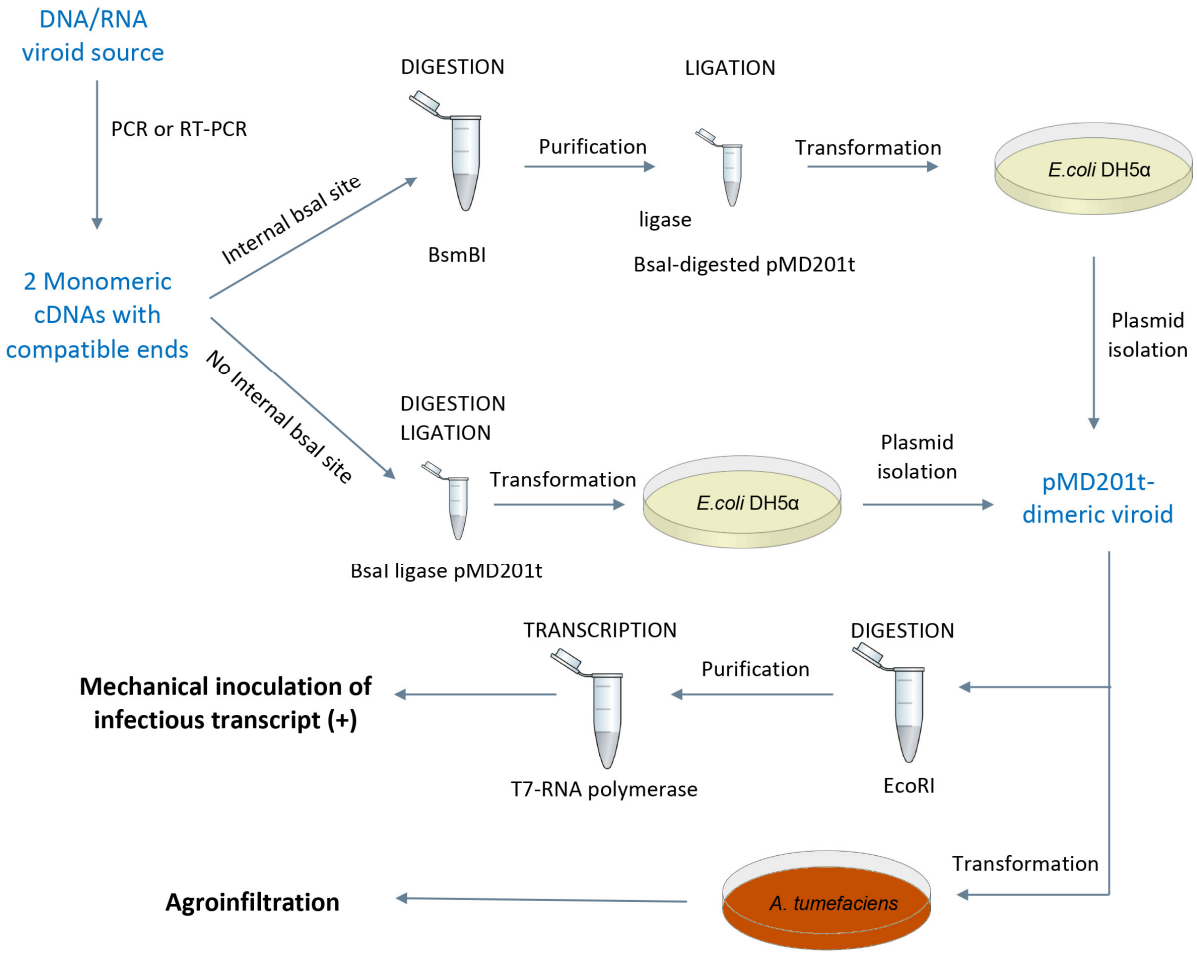


Figure 1.2. Workflow proposed to obtain infectious clones of a viroid. Viroid sequence can be amplified from infected tissue by RT-PCR or from a DNA source by PCR. If the viroid sequence does not contain a BsaI recognition site, the viroidal cDNA can be directly assembled into the binary vector (pMD201t), replacing a lethal gene, in a simultaneous BsaI restriction and ligation. Conversely, if there the viroid contains a BsaI recognition site, it can be cloned using another IIs enzyme, such as BsmBI. Once digested and purified, the viroid cDNA is dimerized by ligation to a previously digested pMD201t. The dimeric viroidal cDNA cloned into pMD201t (pMD201t-viroid) can be used to generate the infectious RNA transcript *in vitro*, using T7 RNA polymerase onto a linearized plasmid (digested with EcoRI). Additionally, the pMD201t-viroid can be transformed into *Agrobacterium tumefaciens* for transient plant transformation and subsequent production of infective RNA *in vivo*.

The dimeric viroid-derived clones are highly infectious

Once the constructs were obtained and sequenced, agro-mediated inoculation bioassays were conducted to analyze the infectivity of the HSVd- and ELVd- derived clones. Cotyledons of ten cucumber plants were agroinfiltrated with pMD201t-HSVd as this wide-host range viroid produces characteristic symptoms in this experimental host (Gómez and Pallás, 2004). All inoculated plants (10/10) were positive for HSVd detection (Figure 1.3A left) and presented the characteristic symptoms of infection at 28 dpi, mainly characterized by growth delay and leaf dwarfism (Figure 1.3C left). A northern blot was carried out to detect the mature forms of HSVd-RNA in systemic leaves of three representative infected-plants (Figure 1.3B left). Related to ELVd-derived clone

assay, eggplant cotyledons were agroinfiltrated with pMD201t-ELVd. In coincidence with the observed in HSVd-infection, hybridization assays evidenced that ELVd could be efficiently detected by dot blot in all inoculated plants at 28 dpi (10/10) (Figure 1.3A right). Similarly, mature forms of ELVd were detected by northern blot in systemic leaves recovered from three representative infected- plants (Figure 1.3B). As was previously described [29], ELVd infects eggplant without producing any visible symptom (Figure 1.3C right).

Finally, to check the efficiency of the constructed dimeric clone as template for *in vitro* transcription assays (+) strand transcripts of the HSVd and ELVd were obtained using T7 RNA polymerase (Supplementary Figure 1.3).

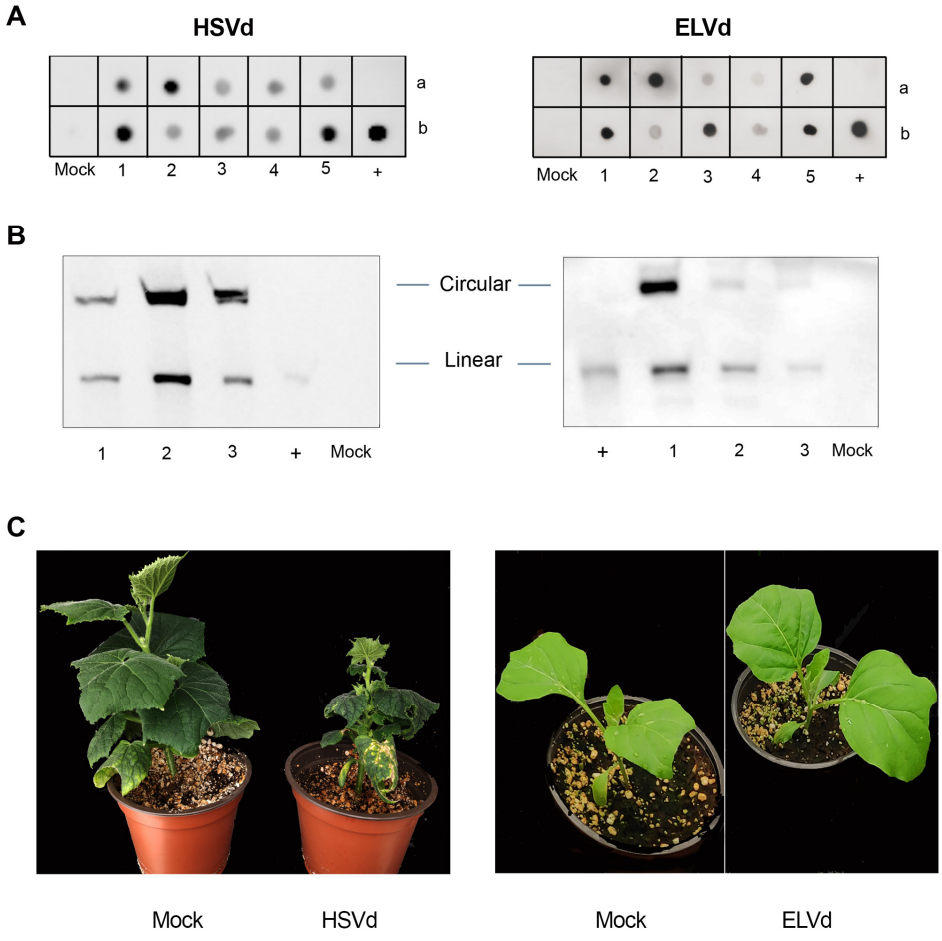


Figure 1.3. Infectivity of viroid constructs. A) Dot blot of systemic leaves at 21 dpi of ten plants (a1-5/b1-5) agroinfiltrated with HSVd construct (left panel) and ELVd construct (right panel) or and two plants agroinfiltrated with empty vector (Mock a/b). HSVd and ELVd transcripts of plus polarity were used as positive control (+b). **B**) Northern blot of systemic leaves at 28 dpi of representative cucumber and eggplant plants agroinfiltrated with HSVd construct (1-3 left) and with ELVd construct (1-3 right) respectively. In both case, plants inoculated with empty vectors were used as mock control. Monomeric linear transcript of HSVd and ELVd, respectively were used as positive controls (+). **C**) Figure showing representative symptomatic (cucumber-HSVd) and asymptomatic (eggplant-ELVd) infected plants at 28 dpi.

Discussion

We have developed an optimized strategy for oriented assembly of viroid-derived cDNA to generate dimeric infectious clones. The process consists of direct cloning into a binary vector that can be transformed into *Agrobacterium tumefaciens* to establish viroid infection by agroinfiltration or employed to generate RNA transcripts *in vitro* using T7 RNA polymerase.

Type IIs restriction enzymes BsaI and BsmBI were chosen because they cleave outside their recognition site at N1/N5 and thus these four nucleotides can be freely selected, being arranged in our strategy to produce a seamless fusion between the monomeric viroid cDNAs and to the binary vector pMD201t. Other IIs enzymes whose cleavage products result in four nucleotides overhangs exist, such as BtgZ1 (N10/N14), BveI (N4/N8) and BbS1 (N2/N6) but are less desirable because of the larger separation between the recognition and cleavage sites, requiring longer primers. In particular, the binary vector pMD201t was designed with BsaI recognition sites and we propose the use of those for any compatible sequence because it is considerably more affordable than the other IIs restrictions enzymes; in fact, most of the type IIs-based assembly methods rely on this enzyme (Lin and O'Callaghan, 2018).

Considering the small size of viroids genome (246-434 nucleotides), the occurrence of both restriction sites (BsaI and BsmBI) in the same viroid is extremely unlikely -lower than 0,20% of the know viroid variants- (Supplementary Table 1.2), but in any case, could be overcome using the other aforementioned IIs restriction enzymes or by the assembly of the fractioned monomers produced by internal IIs enzyme cleavage. This later strategy would be feasible unless the four nucleotides left overhangs were compatible with the four chosen for dimerization or vector assembly, which is a very remote possibility. However, that approach will decrease the efficiency of ligation as the number of parts to be ligated is increased. These considerations clearly validate the universality of our strategy for constructing infectious clones of any viroid. Additionally, this strategy could be applied to expand the possibilities of site directed mutagenesis in viroids, by using mutagenic primers to amplify the monomeric sequence, which could be subsequently dimerized, therefore avoiding the apparition of unexpected mutations previously reported when direct mutagenesis of the dimeric viroid is used (Sanjuán and Daròs, 2007).

Nonetheless, we have developed this assembly strategy for obtaining dimers of viroids, but it can be repurposed for oriented seamless ligation of any sequence directly onto binary vectors with an optimal efficiency. Consequently, our approach illustrates the applicability of type IIs restriction enzymes and the lethal gene *ccd*B to design efficient and affordable cloning approaches into binary vectors.

In conclusion, the purpose of this study was to develop a global (applicable to phylogenetically unrelated families), innovative, and quickly strategy to construct infectious viroid clones suitable to be employed in *Agrobacterium*-mediated inoculation and/or as template for *in vitro* transcription of viroid RNA. Our results support, that this methodology constitutes a valuable tool for viroid research and reinforce the applicability of type IIs restriction enzymes and the lethal gene *ccdB* to design efficient and direct cloning approaches of PCR products into binary vectors.

Material and methods

Binary vector construction

A binary vector suitable for the direct cloning and transcription/expression of dimeric viroid was engineered from the binary vector pMDC32B-AtMIR390a-Bc (Carbonell et al., 2014). This vector was digested with restriction enzymes EcoRI and Hind III (Thermo Scientific™, Waltham, MA, USA), and the resulting fragment of 8.9 kb was excised from a 1% agarose gel and purified using GeneJET Gel Extraction Kit (Thermo Scientific™, Waltham, MA, USA). All reactions were performed following manufacturer's instructions. DNA fragments were amplified with PrimeSTAR™ HS DNA Polymerase (Takara, Kusatsu, Japan) and ligations were set with an insert:vector ratio of 3:1 and 3U of T4 DNA ligase (Promega, Madison, WI, USA) being transformed into DB3.1 *Escherichia coli* cells. In order to obtain a positive selection vector, the lethal gene *ccd*B was cloned into an engineered pSK vector designed for cloning, after digestion with NcoI and NheI, a sequence between a duplicated promoter CaMV 35S and the Potato Protease inhibitor II terminator (PoPit)(Genoves et al., 2006). This vector was modified to eliminate a HindIII recognition after the PoPit terminator, by amplifying the plasmid with inverted oligonucleotides: Fw mut-HindIII and Rv mut-HindIII. The lethal gene *ccd*B was amplified from pMDC32B-AtMIR390a-Bc using Fw ccdB-NcoI/b and Rv ccdB-NheI/b (designed to generate compatible ends with NcoI and NheI) and ligated to the aforementioned pSK vector after BsaI digestion (Thermo Scientific™, Waltham, MA, USA). Additionally, 2X35S:*ccd*B:PoPit cassette was amplified with Fw T7-35S HindIII, to introduce the T7 RNA polymerase promoter, and Rv M13, digested with EcoRI and HindIII (Thermo Scientific™, Waltham, MA, USA) and ligated to the above mentioned backbone of pMDC32. The resulting vector was named pMD201t, number 201 indicates 2X35S:*ccd*B:PoPit, respectively, and the letter t makes references to the T7 promoter. All oligonucleotides used for generating this construct are listed in Supplemental Table 1.1.

Dimeric viroid cDNA construction

Monomeric form of *Hop stunt viroid* (Y09352.1), previously cloned in a pBluescript II SK vector (Gómez and Pallás, 2004), was used as a template to generate the two DNA fragments required for the dimer assembly. PCR reactions were performed with the following reaction mixture: 1.25U PrimeSTAR™ HS DNA Polymerase (Takara, Kusatsu, Japan), 5 µL of 5X Buffer, 2 µL of 2.5 mM dNTP mixture, 25 µM of each primer, 75 ng of plasmid template and sterilized water up to 25 µl. PCR conditions were 30 cycles of 10 s at 98°C, 5 s at 55°C and 21 s at 72°C. One reaction was performed with the primers Fw D1-HSVd and Rv D2-HSVd, and the other with Fw D3-HSVd and Rv D4-HSVd (Supplemental Table 1.1). These DNA fragments were purified together using

GeneJET Gel Extraction Kit (Thermo Scientific™, Waltham, MA, USA) and a one-pot reaction was set as follows: 10 U of BsaI (Thermo Scientific™, Waltham, MA, USA), 3U T4 DNA ligase (Promega, Madison, WI, USA), 1µL of ligase buffer 10X (Promega, Madison, Wi, USA), 50 ng of pMD201t and 300 ng of the digested DNA fragments in a final volume of 10 µL. The incubation was performed using a thermocycler with the following conditions: an initial step of 20 minutes at 37 °C, 20 cycles of 1 minutes at 37 °C and 4 minutes a 16 °C, finally holding on the temperature at 16 °C until transformation.

Monomeric form of ELVd cDNA (AJ536613) was synthetized as gBlocks® (Integrated DNA Technologies Inc., Coralvelle, IA, USA) and used as a template, as described above, for PCR amplification with Fw D1-ELVd Rv D2-ELVd and Fw D3-ELVd Rv D4-ELVd. PCR conditions were identical to those aforementioned except that 50 ng of the gene fragment was used for each reaction as template. PCR products were purified together using GeneJET Gel Extraction Kit (Thermo Scientific™, Waltham, MA, USA) and digested with 10 U of BsmBI (Thermo Scientific™, Waltham, MA, USA) according to manufacturer instructions in a final volume of 40 µl, and eventually purified and concentrated into a final volume of 20 µl using GeneJET Gel Extraction Kit (Thermo Scientific™, Waltham, MA, USA). Finally, a ligation reaction was set with: 3U T4 DNA ligase (Promega, Madison, Wi, USA), 1µL of ligase buffer 10X (Promega, Madison, WI, USA), 50 ng of BsaI digested pMD201t and 300 ng of the digested monomer fragments in a final volume of 10 µL. This ligation was incubated for 2 hours at room temperature.

1-2 µL of the HSVd/ELVd cDNAs ligation to pMD201t, respectively, was transformed into DH5-Alpha electro competent cells and plated onto kanamycin agar plates (Supplementary Figure 1.1). 8 colonies were analyzed by PCR colony using Go-Taq (Promega, Madison, WI, USA) and oligonucleotides Fw 35S-AMV and Rv Popit (Supplementary Figure 1.2). Plasmid extraction was performed with GeneJET Plasmid Miniprep Kit (Thermo Scientific™, Waltham, MA, USA) and resulting constructs were sequenced using Rv Popit. The oligonucleotides used are listed in Supplementary Table 1.1.

Viroid inoculation

Cotyledons of *Cucumis sativus* cv. Marketer and *Solanum melongena* cv. Black Beauty were agro-infiltrated with a culture of *A. tumefaciens* strain C58 harbouring the correspondent binary vector, pMD201t-HSVd for *C. sativus* and pMD201t-ELVd for *S. melongena*. The overnight grown bacterial culture was diluted in infiltration buffer (MES 0.1 M, $MgCl_2$ 1 M) up to an optical density at 600 nm of 1 and injected on the abaxial side of one cotyledon using a needle-less syringe. Plants were kept in a photoperiod of 16 hours under visible light and 30 °C (light)/25 °C (darkness) for *C. sativus* and 25 °C (light) /18 °C (darkness) for *S. melongena*. Samples of systemic leaf tissue were collected at 21 and 28 days after viroid inoculation.

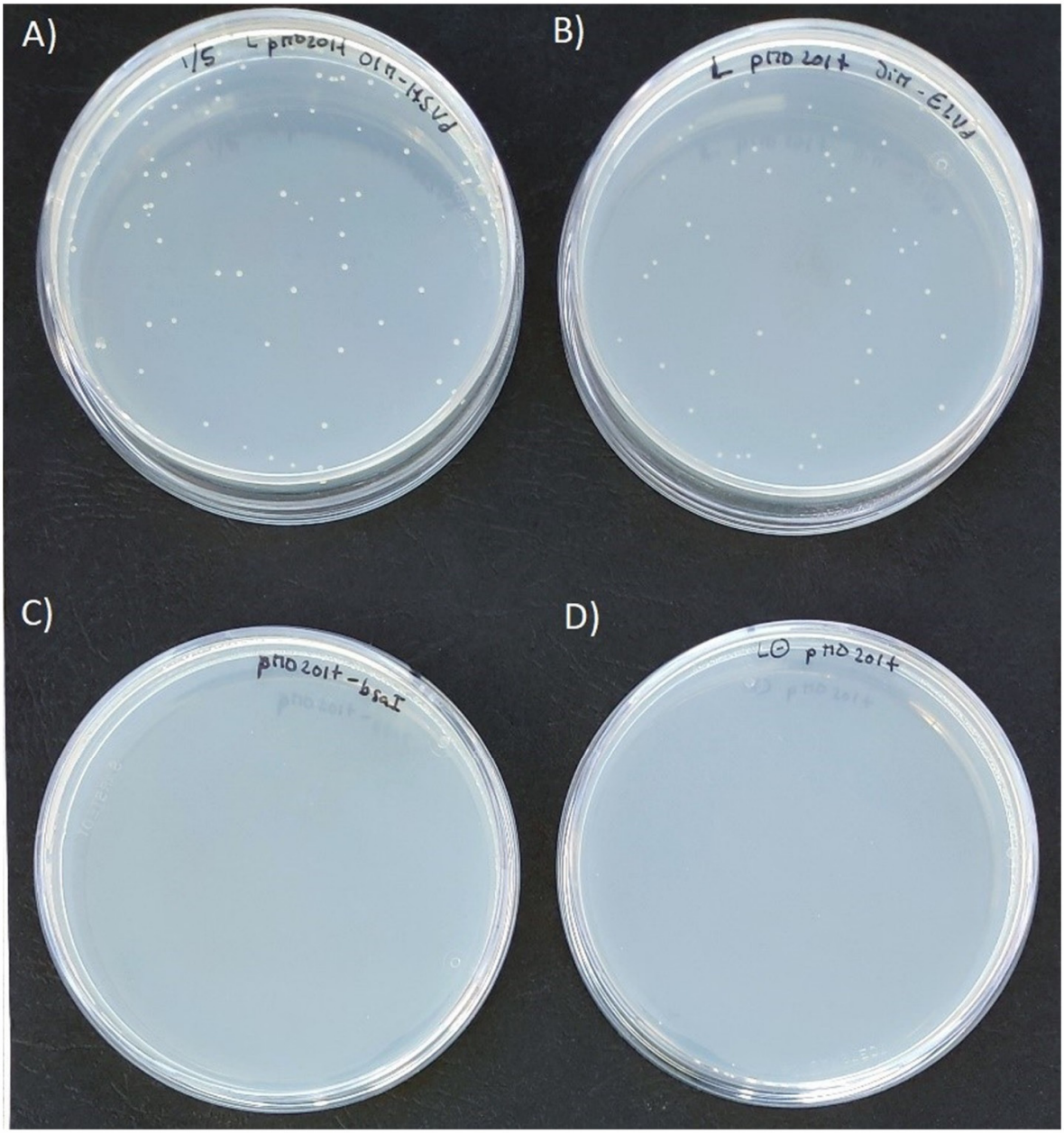
***In vitro* transcription**

Viroid transcripts were generated by transcription of 400 ng of linearized pMD201t HSVd/ELVd (Digested with EcoRI) with T7 RNA polymerase (Takara, Kusatsu, Japan) for 3 hours according to manufacturer instructions. 1 µL of each 10 µL reaction was loaded into a sterile 1% agarose gel, by serial dilutions (0.1, 0.3 and 0.6 µL). RiboRuler High Range RNA Ladder (Thermo Scientific™, Waltham, MA, USA) was used to estimate the RNA concentration; as for the loaded volume (0.83 µL) each ladder band corresponds to 50 ng.

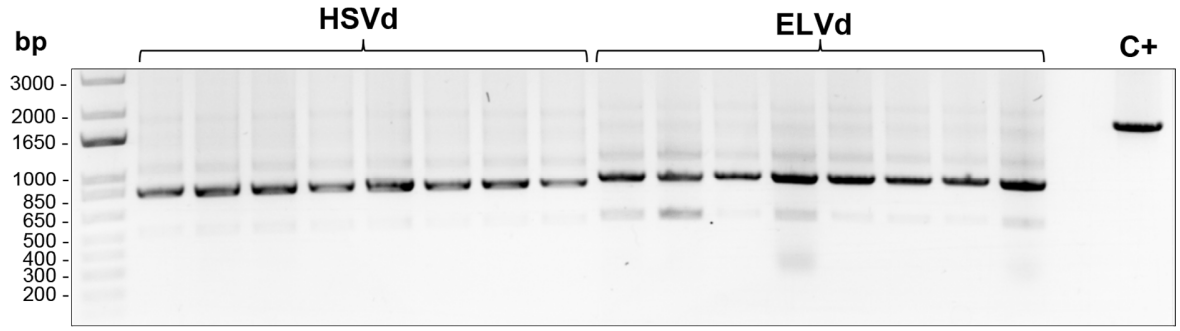
RNA extraction and northern/dot blot

Total RNA was extracted from systemic leaves as described previously (Genoves et al., 2006). 2-5 µg of total RNA per sample were mixed with solid urea and then loaded into a PAGE 5% UREA 8M and TBE 89 mM gel. RNA electrophoresis was performed at 200 V for one hour and then RNA was transferred to a nylon membrane using MiniProtean 3 system (BioRad, Hercules, CA, USA). Transference conditions were 100 V for 1 hour at 4 °C in TBE buffer 1X. Nucleic acids transferred to the membrane (northern) or directly applied onto the nylon membrane (dot, 1 µL of total RNA per sample) were covalently fixed by using ultraviolet light (700x100 J/cm^2). Hybridisation and chemiluminescent detection were performed as previously described (Herranz et al., 2005).

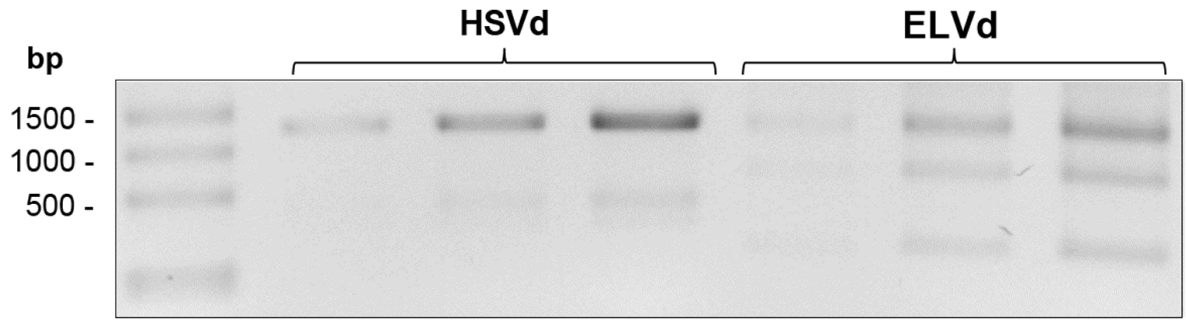
Supplementary material



Supplementary Figure 1.1. Process of cloning of viroid cDNA sequences into pMD201t2 A) Restriction-ligation of HSVd cDNA fragments to pMD 201t. 1 µL of the reaction was transformed into DH5-Alpha electrocompetent cells and 1/5 was plated onto kanamycin agar plates. **B**) Ligation of ELVd cDNA fragments to pMD201t. 2 µL of the ligation were transformed into DH5-Alpha electro competent cells and plated onto kanamycin agar plates. **C**) Control of BsaI-digested pMD201t. 50 ng of pMD201t digested with BsaI were transformed into DH5-Alpha electro competent cells and plated onto kanamycin agar plates. **D**) Control of restriction-ligation of pMD201t. 1 µL of the reaction was transformed into DH5-Alpha electro competent cells and plated onto kanamycin agar plates.



Supplementary Figure 1.2. Colony PCR of pMD201t. Agarose gel 1 % showing PCR amplification of eight individual colonies of pMD201t ligation to HSVd fragments (left) or ELVd fragments (right). The observed size corresponds to the viroid dimer plus part of the terminator and promoter sequences. A positive control of PCR amplification (C+) was the undigested vector.



Supplementary Figure 1.3. *In vitro* transcription of pMD201t. 400 ng of linearized pMD201t HSVd/ELVd was transcribed with T7 RNA polymerase. Serial dilutions of linearized pMD201-HSVd transcription (left) and linearized pMD201-ELVd (right). RiboRuler High Range RNA Ladder (Thermo Fisher Scientific) was used to estimate the RNA concentration, as for the loaded volume (0.83 µL) each ladder band corresponds to 50 ng.

Supplementary Table 1.1

Oligonucleotide	Sequence 5'-3'
Fw D1-HSVd	GGTCTCCCATGGCTGGGGAATTCTCGAGTTG
Rv D2-HSVd	GGTCTCAAGGGGCTCAAGAGAGGAT
Fw D3-HSVd	GGTCTCCCCCTGCTGGGGAATTCTCGAGTTG
Rv D4-HSVd	GGTCTCTCTAGAAGGGGCTCAAGAGAGGAT
Fw D1-ELVd	GACGTCTCCCATGGGGTGGTGTGTGCCACCCCT
Rv D2-ELVd	GCGCCGTCTCCTATGGGGAGAGGTCGTCCTCTATC
Fw D3-ELVd	GCAGCGTCTCTCATAGGGTGGTGTGTGCCACCCCTG
Rv D4-ELVd	GCGCCGTCTCGCTAGCTATGGGGAGAGGTCGTCCTCTAT
Rv Popit	TGGATGATCTCTTTCTCTTATTCAG
Fw 35S-AMV	CATTTGGAGAGGGTTTTTATTTTT
Fw ccdB-NcoI/b	GCACCATGGGAGACCATTAGGCACCCCAGG
Rv ccdB-NheI/b	GCTGCTAGCGAGACCGTCGAGGTGCAGAC
Fw mut-HindIII	CCATATACAAGCcTTGATAATCGAATTCC
Rv mut-HindIII	GGAATTCGATTATCAAgGCTTGTATATGG
Fw T7-35S HindIII	CGAAAGCTTTAATACGACTCACTATAGGCTTACTCGAGCTCGGGCCCTCTACTCC
Rv M13	AGCGGATAACAATTTCACACAGG

Supplementary Table 1.2. Details of the viroid related sequences containing the restriction sites used in our method. As can be observed, only the 0,19% of the known viroid variants in both *Pospiviroidae* and *Avsunviroidae* families contain both restriction sites. Viroid sequences were recovered from https://www.ncbi.nlm.nih.gov/nuccore.

	Number	**Percentage**
Viroid related sequences	11395	100%
Containing BsaI sites	645	5.66%
Containing BsmBI sites	503	4.41%
Containing both BsaI and BsmBI sites	22	0.19%

CHAPTER 2

Might exogenous circular RNAs act as protein-coding transcripts in plants?

Joan Marquez-Molins[1,2], José Antonio Navarro[2], Luis Cervera Seco[1], Vicente Pallas[2] and Gustavo Gomez[1*].

[1] Institute for Integrative Systems Biology (I2SysBio), Consejo Superior de Investigaciones Científicas (CSIC) - Universitat de València (UV), Parc Científic, Cat. Agustín Escardino 9, 46980 Paterna, Spain.

[2] Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de Investigaciones Científicas (CSIC) - Universitat Politècnica de València, CPI 8E, Av. de los Naranjos s/n, 46022 Valencia, Spain.

[*] *Corresponding author*: E-mail: gustavo.gomez@csic.es

Published as RNA biology 2021, 18, 98-107, doi: 10.1080/15476286.2021.1962670

PERSONAL CONTRIBUTION

I performed all the experiments, analyzed and discussed the results, and performed computational analysis.

ACKNOWLEDGMENTS

The authors thank Dr. J.A. Daròs (IBMCP) for kindly providing the plasmids containing PSTVd-RG1 cDNA and the original ELVd sequence; and also thank Dr. M. De la Peña for assistance with the prediction of RNA tertiary structure. This work was supported by the Spanish Ministry of Economy and Competitiveness (co-supported by FEDER) Grants BIO2017-88321-R (VP) and AGL2016-79825-R (GG). I have been recipient of a predoctoral contract from the ACIF programme (ACIF-2017-114) of the Conselleria d'Educació, Investigació, Cultura i Esport Generalitat Valenciana.

Abstract and keywords

ABSTRACT

Circular RNAs (circRNAs) are regulatory molecules involved in the modulation of gene expression. Although originally assumed as non-coding RNAs, recent studies have evidenced that animal circRNAs can act as translatable transcripts. The study of plant-circRNAs is incipient, and no autonomous coding plant-circRNA has been described yet. Viroids are the smallest plant-pathogenic circRNAs known to date. Since their discovery 50 years ago, viroids have been considered valuable systems for the study of the structure-function relationships in RNA, essentially because they have not been shown to have coding capacity. We used two pathogenic circRNAs (*Hop stunt viroid* and *Eggplant latent viroid*) as experimental tools to explore the coding potential of plant-circRNAs. Our work supports that the analyzed viroids contain putative ORFs able to encode peptides carrying subcellular localization signals coincident with the corresponding replication-specific organelle. Bioassays in well-established hosts revealed that mutations in these ORFs diminish their biological efficiency. Interestingly, circular forms of HSVd and ELVd were found to co-sediment with polysomes, revealing their physical interaction with the translational machinery of the plant cell. Based on this evidence we hypothesize about the possibility that plant circRNAs in general, and viroids in particular, can act, under certain cellular conditions, as non-canonical translatable transcripts.

KEYWORDS

Circular RNAs, Plant coding circRNAs, non-canonical transcripts, Plant pathogenic RNAs, viroids, viroid-derived peptides.

Introduction

Circular RNAs (circRNAs) are covalently closed transcripts found in diverse organisms that play important regulatory roles (Ji et al., 2019). According to its origin, circRNAs can be classified as endogenous (*en*-circRNAs), originated from the transcription of the own genome or exogenous (*ex*-circRNAs), associated with external agents (Patop et al., 2019). Viroids (a class of ex-circRNAs) were discovered as replicating, low molecular weight infectious RNAs (Diener, 1971) with circular nature (Sanger et al., 1976). Viroids, thus, can be considered as the first circRNAs found in any organism. A few years later, the presence of virus-derived circRNAs in the cytoplasm of human cells was reported (Hsu and Coca-Prados, 1979). Endogenous circRNAs were first found in humans associated with the non-canonical splicing of cancer-related transcripts (Nigro et al., 1991). Diverse experimental evidence has revealed that multiple types of *en*-circRNAs are expressed in animals (Memczak et al., 2013). The first global identification of *en*-circRNAs in plants was done in rice and arabidopsis (Ye et al., 2015). More recent studies have described the existence of *en*-circRNAs in diverse plant species (Darbani et al., 2016; Tan et al., 2017; Wang et al., 2017a; Wang et al., 2017b; Chen et al., 2018a; Zhu et al., 2019; Han et al., 2020). The majority of human *en*-circRNAs are a few hundred nucleotides in length (Patop et al., 2019), while soybean *en*-circRNAs are mainly between 150-600 bp (Zhao et al., 2017). A common characteristic of both animal and plant *en*-circRNAs is that they are dependent on RNA polymerase II-mediated transcription and non-canonical RNA processing (Zhao et al., 2019).

Under a functional viewpoint, animal *en*-circRNAs act predominantly as miRNA sponges (Hansen et al., 2013). Moreover, it has been proposed that they may also be involved in modulating the transcriptional activity (Li et al., 2015), and in the memory of the transcriptional history of the cell (Patop et al., 2019). The potential role of plant circRNA remains largely unknown as only two *en*-circRNAs have been functionally characterized (Conn et al., 2017; Fan et al., 2020). Plant *en*-circRNAs possess features that differ from those described in animals, like a lower prediction to be potential miRNA sponges (Ye et al., 2015). However, their functional characterization could leverage animal circRNA studies (Lai et al., 2018).

In general, circRNAs have been considered as non-coding RNAs (ncRNAs) (Guo et al., 2014). However, recent evidence supports that certain animal *en*-circRNAs are translated *in vivo* (Patop et al., 2019) suggesting that the coding potential of these assumed ncRNAs has been underestimated and that may exert more biological functions than previously predicted. To date, no plant *en*-circRNAs has been reported to generate proteins and only non-autonomous coding activity has been reported for a virus-satellite circRNA in rice (AbouHaidar et al., 2014).

Nonetheless, the coding potential of plant circRNAs has not been deeply investigated. Elucidating the functional roles of circRNAs, in translation, might constitute a primary research topic in plant circRNAs (Zhao et al., 2019).

Viroids are thus, naked *ex*-circRNAs, without apparent coding capacity, that infect host-plants causing phenotypic effects ranging from severe symptoms to latent infections (Flores et al., 2015). In coincidence with *en*-circRNAs, viroid replication is dependent on plant-endogenous RNA polymerases, and their genome (≅ 300 nt) is comparable to the predominant size of *en*-circRNAs. Additionally, these autonomous pathogenic RNAs also exhibit differential subcellular compartmentalization; as members of the *Pospiviroidae* family replicate and accumulate in the nucleus, while the *Avsunviroidae* develop their life cycle in chloroplasts (Di Serio et al., 2014). Due to their particular biological characteristics, viroids have emerged as a valuable tool to explore the regulatory pathways mediated by RNA in plants (Gómez and Pallás, 2013; Steger and Riesner, 2018). Prompted by the unexpected discovery of the coding activity of circRNAs in animals, here we use these pathogenic *ex*-circRNAs as an experimental model to explore the potential of plant circRNAs to encode functional peptides. In this opinion article we hypothesize about the possibility that under specific (and today unknown yet) cellular conditions, *ex*-circRNAs could act as protein-coding transcripts. We analyzed representative members of the two viroid families replicating in either nucleus or chloroplasts and showed that HSVd and ELVd circRNAs contain conserved ORFs able to encode functional peptides carrying specific subcellular localization signals that direct fluorescent-fused proteins to the corresponding organelle where replication/accumulation takes place. Through bioassays in well-established hosts, we show that mutations in these ORFs diminish the biological efficiency of these pathogenic *ex*-circRNAs. Finally, circular forms of both, HSVd and ELVd RNA were found to co-sediment with plant polysomes *in vivo,* revealing their physical interaction with the translational machinery of the plant cell.

Results

HSVd and ELVd contains putative ORFs able to encode peptides

In order to analyze their potential coding capability, we computationally inferred all the possible peptides derived from the *plus* strand of infectious cDNA clones (Marquez-Molins et al., 2019) of *Hop stunt viroid* (HSVd) and *Eggplant latent viroid* (ELVd), members of families *Pospiviroidae* and *Avsunviroidae*, respectively. In viroids, the mature RNA form predominantly accumulated in the host is referred to as the plus strand. HSVd has been found inducing symptomatic infection in a wide range of hosts and has been frequently used to explore diverse aspects of viroid-host interactions (Martinez et al., 2014; Castellano et al., 2016b; Wei et al., 2019; Marquez-Molins et al., 2021). ELVd induces asymptomatic infections and is assumed as a valuable model to study molecular aspects of the family *Avsunviroidae* (Daròs, 2016).

Viroid sequences were translated into the three possible frames using any codon as start. Only peptides longer than 45 amino acids (according to the criteria explained in experimental procedures) derived from each ORF were selected for further analysis (Figure 2.1 and Supplementary Figure 2.1A and B). The putative HSVd-ORFs 1 and 2 (H-ORF1 and H-ORF2) were 48 and 98 amino acids in length, respectively. The third ORF (H-ORF3) constitutes a unique reading frame without any stop codon (Figure 2.1A). ELVd-ORFs were 110 (E-ORF1), 87 (E-ORF2), and 59 (E-ORF-3) amino acids in length (Figure 2.1B).

To determine if these ORFs are conserved in HSVd and ELVd, the predicted ORFs were aligned against the ORFs identified in the complete sequences publicly available of HSVd (779) and ELVd (104). The conservation rate was estimated considering the presence (in percentage) of the predicted ORFs in all analyzed accessions with identity value (between potential peptides) higher than 50%. Conservation rates lower than 50% were considered as not significant. As it is shown in Figure 2.1C, H-ORF3 was detected in 59% of the analyzed sequences with an identity value mean of 68.12%. The highest conservation value observed in ELVd was for the E-ORF1, which presents an identity mean of 81.17% in 96.6% of the accessions. The other predicted ORFs (E-ORF2 and E-ORF-3) present identity means of 75.93% and 65.31% in the 63.46% and 73.07% of the analyzed accessions, respectively (Figure 2.1D). The consensus sequences of the most conserved ORFs detected in HSVd and ELVd circular forms are shown in the Figure 2.1E and 2.1F, respectively.

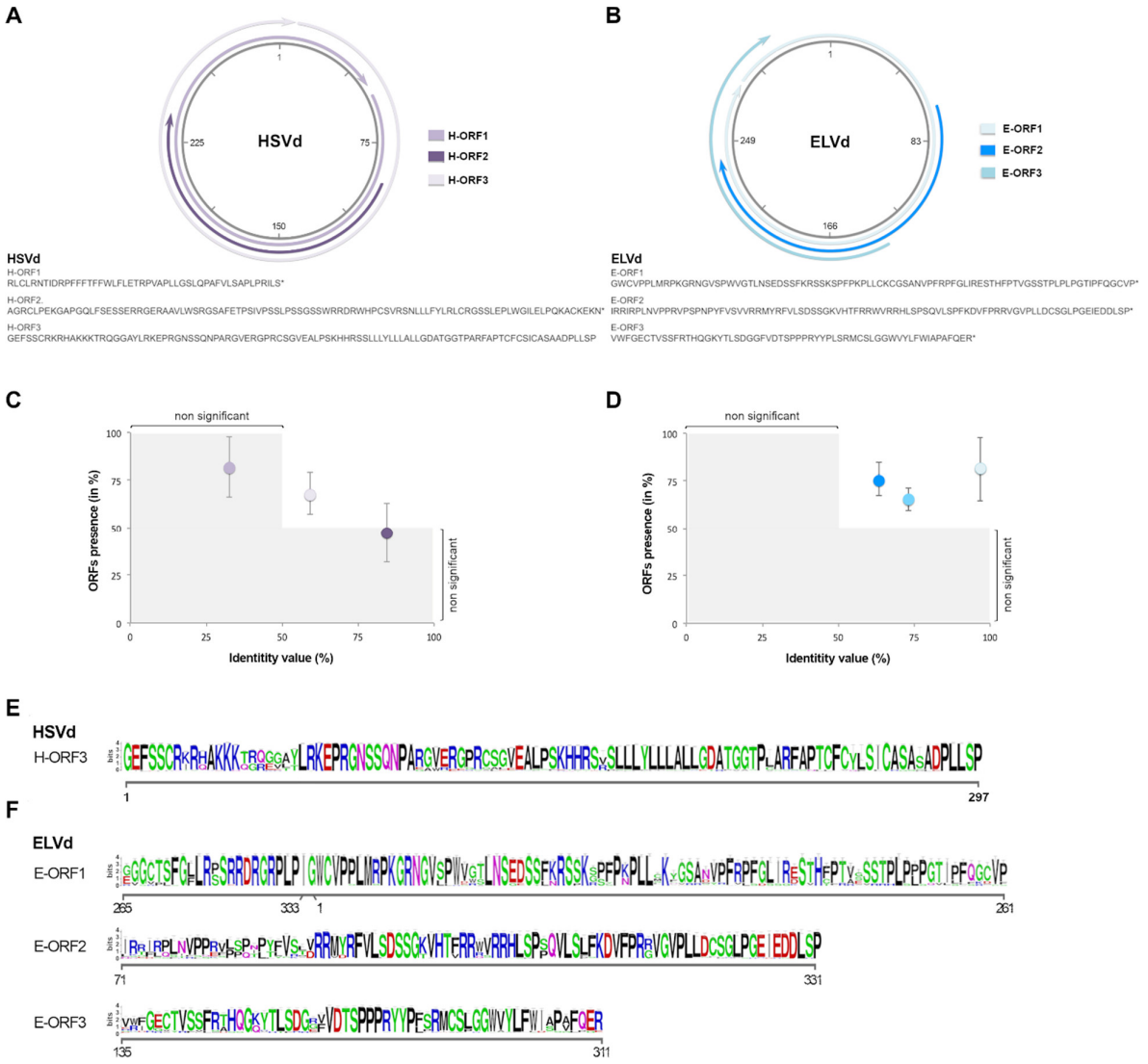


Figure 2.1. Computational prediction of putative ORFs in *ex*-circRNAs. Detailed location (upper panel) and peptide sequences (lower) of the ORFs identified in HSVd (**A**) and ELVd (**B**) infectious clones. Graphic representation of the conservation rates estimated for the ORFs inferred in HSVd (**C**) and ELVd (**D**) accessions recovered from public repository. Consensus sequence of the conserved peptides predicted to be encoded by HSVd (**E**) and ELVd (**F**). Error bars represent the standard error values.

Peptides derived from conserved ORFs (H-ORF3, E-ORF1 and E-ORF2) were blasted against the non-redundant protein database of higher plants (taxid: 3193) at the NCBI. Significant similarities (>50%) were not found for any of the putative peptides. Only a region (51 amino acids) of the E-ORF2 showed a slight identity (31%) with Choline monooxygenase (Supplementary Figure 2.1C), a nucleus-encoded ferredoxin-dependent enzyme located in the chloroplast stroma (Rathinasabapathi et al., 1997).

Putative viroid-derived peptides show specific cellular compartmentalization

To analyze their potential activity, we developed reporter constructs containing the ORFs derived from HSVd and ELVd fused in frame to the amino end of the green (for H-ORFs) and yellow (for E-ORFs) fluorescent proteins cDNAs (Supplementary Figure 2.2). The chimeric cDNAs (H-ORFs/GFP and E-ORFs/YFP) were cloned into a binary vector and transiently expressed in *N. benthamiana* plants.

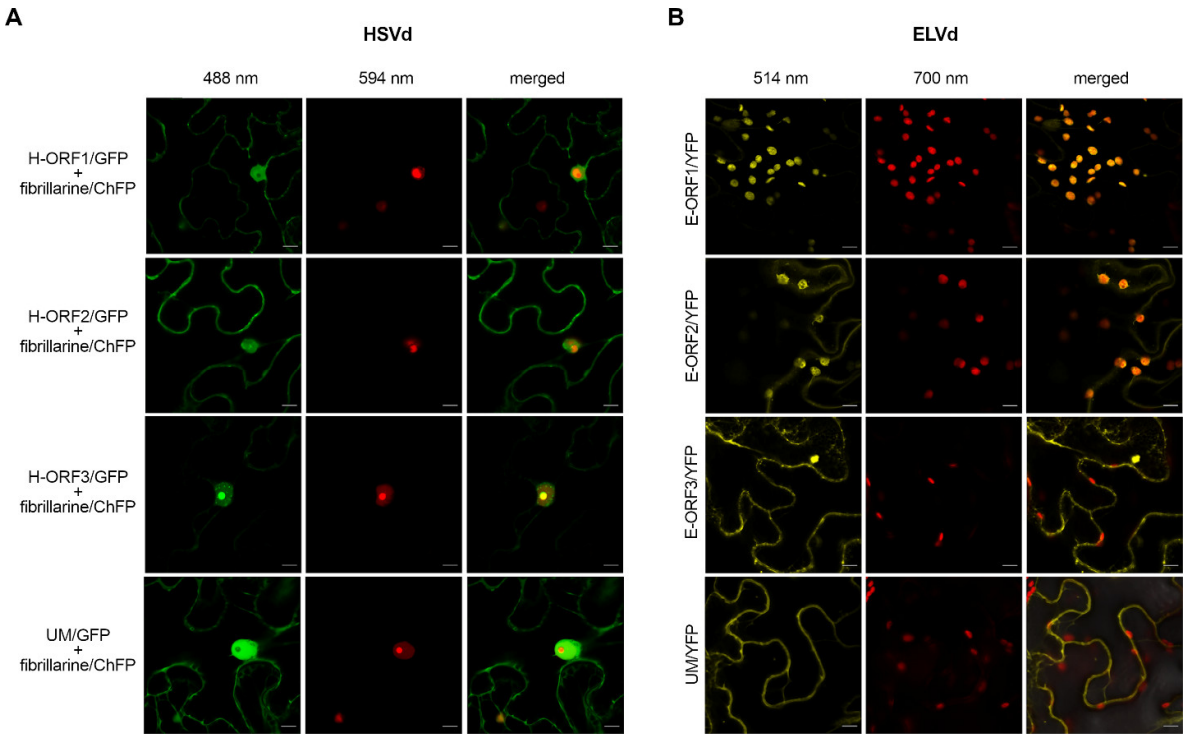


Figure 2.2. Subcellular localization of the putative *ex*-circRNAs-derived peptides. Confocal microscopy images of *N. benthamiana* leaves agro-infiltrated with constructs expressing H-ORFs fused to GFP (**A**) and E-ORFs in frame with YFP (**B**). Unmodified GFP/YFP (UM/GFP and UM/YFP, respectively) were used as control. The fusion fibrillarin-ChFP (central panels in A) was used as a nucleolus marker. Chloroplasts were determined by the chlorophyll auto-fluorescence recovered above 700 nm (central panels in B). Scale bars are 10 μm.

Observation of infiltrated leaves by confocal microscopy revealed that H-ORF1/GFP and H-ORF2/GFP show a cellular localization pattern similar to GFP (Figure 2.2A). Conversely, and in coincidence with a previous observation (Gómez and Pallás, 2007b), H-ORF3/GFP was predominantly accumulated in the nucleus of the analyzed cells (Figure 2.2A). Specifically, the fluorescence signal co-localized with fibrillarin, a well-established marker of the nucleolus (Kalinina et al., 2018). Searching for potential domains responsible for this specific subcellular localization revealed the presence, in the peptide H-ORF3, of a region enriched in basic amino acids conserved in a high proportion of the HSVd sequences described to date (Supplementary Figure 2.3A - left panel). This type of R/K motifs are involved in the specific targeting of peptides to the plant cell nucleolus, acting as *Nucleolar Localization Signals* (NoLS) (Carmo-Fonseca et al.,

2000). To evaluate if this putative R/K motif is responsible for the specific subcellular localization of the H-ORF3 peptide, we generated a reporter (H-RK/GFP) carrying this potential domain fused to GFP. Transient expression assays revealed that the protein derived from the H-RK/GFP construct also accumulates in the nucleolus supporting that this H-ORF3 domain acts as a functional NoLS (Supplementary Figure 2.3B). A similar domain enriched in basic amino acids was identified in *Potato spindle tuber viroid* (PSTVd), the type species of the family *Pospiviroidae* (Supplementary Figure 2.3A – right panel). The functionality of this potential NoLS was validated by transient expression (Supplementary Figure 2.3B).

Regarding the putative ELVd-derived peptides, we observed that the chimeric E-ORF1/YFP was predominantly localized in chloroplasts (Figure 2.2B), in agreement with the *in silico* prediction as a chloroplast transit peptide (ChTP) (Supplementary Figure 2.4A). Remarkably, the second ORF identified in the ELVd (E-ORF2) was also predicted to act as a ChTP (Supplementary Figure 2.4A). Transient expression assays revealed that E-ORF2/YFP peptide was also transported to chloroplasts (Figure 2.2B). Finally, E-ORF3/YFP showed a localization pattern comparable to YFP. The vast majority of the nucleus-encoded proteins destined to plastids are targeted into the chloroplast as a precursor whose transport is facilitated by a cleavable N-terminal ChTP (Bruce, 2000). To evaluate if the peptides E-ORF1 and E-ORF2 are transported to chloroplast through this canonical model, chimeric E-ORF1/YFP and E-ORF2/YFP proteins transiently expressed in *N. benthamiana* leaves were analyzed by western blot assays. Our results revealed that both peptides derived from this *ex*-circRNA are cleaved during the transport process to chloroplasts (Supplementary Figure 2.4B). To gain additional functional evidence, we transiently expressed two constructs containing E-ORF1 and E-ORF2 partial peptides lacking the predicted ChTP region fused to YPF. As expected, the YFP fussed to both depleted peptides showed a localization pattern comparable to free-YFP (Supplementary Figure 2.4C).

Furthermore, we also analyzed the subcellular compartmentalization of H-ORF3/GFP and E-ORF1/YFP in *N. benthamiana* and eggplant leaves inoculated with HSVd and ELVd, respectively. The obtained data revealed that the specific accumulation (in nucleolus for H-ORF3/GFP and chloroplasts for E-ORF1/YFP) was also observed in viroid-infected tissues (Supplementary Figure 2.5).

Inactivation of putative ORFs affects biological efficiency of ex-circRNAs

To evaluate if the potential peptides derived from *ex*-circRNAs can exert an effect on the viroid infection cycle, we analyzed the viroid accumulation in cucumber and eggplant leaves infected with mutant HSVd and ELVd variants, respectively. Sequence modifications were designed to truncate the inferred H-ORF3 and E-ORF1 while maintaining (without relevant changes in the free

energy values) the predicted RNA secondary structure of the viroids (Figure 2.3A and B – Supplementary Figures 2.6 and 2.7). Furthermore, we specifically modified positions in which a stop codon could be introduced with a single nucleotide change, and designed control mutants that do not alter the putative ORF but are in the same structural domain.

Two HSVd-derived mutants (H-ORF3/A24U and H-ORF3/A36U) were designed harboring the substitution A/U in the positions 24 and 36, respectively. These modifications introduce two stop codons UAA and UAG in the R/K region responsible for the nucleolar targeting. A third mutant (H-ORF3/C266A) contains a C/A substitution that introduces the stop codon UGA out of the functional NoLS but in the same region of the secondary structure of the viroid. As an additional control we modified the position 266 with a C/U silent substitution that maintains unaltered the H-ORF3 (H-ORF3/C266U) (Figure 2.3A).

To analyze the functionality of the putative peptides derived from ELVd, we constructed mutants carrying a double substitution C/A-G/U in the positions 83 and 95, respectively, which truncates the E-ORF1 with the stop codon UAG (E-ORF1/C83A-G95U) (Figure 2.3B). The mutant E-ORF2/C212U-G317A contains a double substitution that introduces the stop codon UGA in the E-ORF2. Finally, two control mutants (E-ORF1/C83G-G95C and E-ORF2/C212G-G317C) containing substitutions in the same ELVd-sequence positions but preserving the secondary structure and not disrupting any ORF, were also obtained. Analysis of the cleavage efficiency of transcripts derived from the ELVd mutants *in vitro* revealed that the modifications do not affect the autocatalytic activity of the hammerhead ribozymes (Supplementary Figure 2.8).

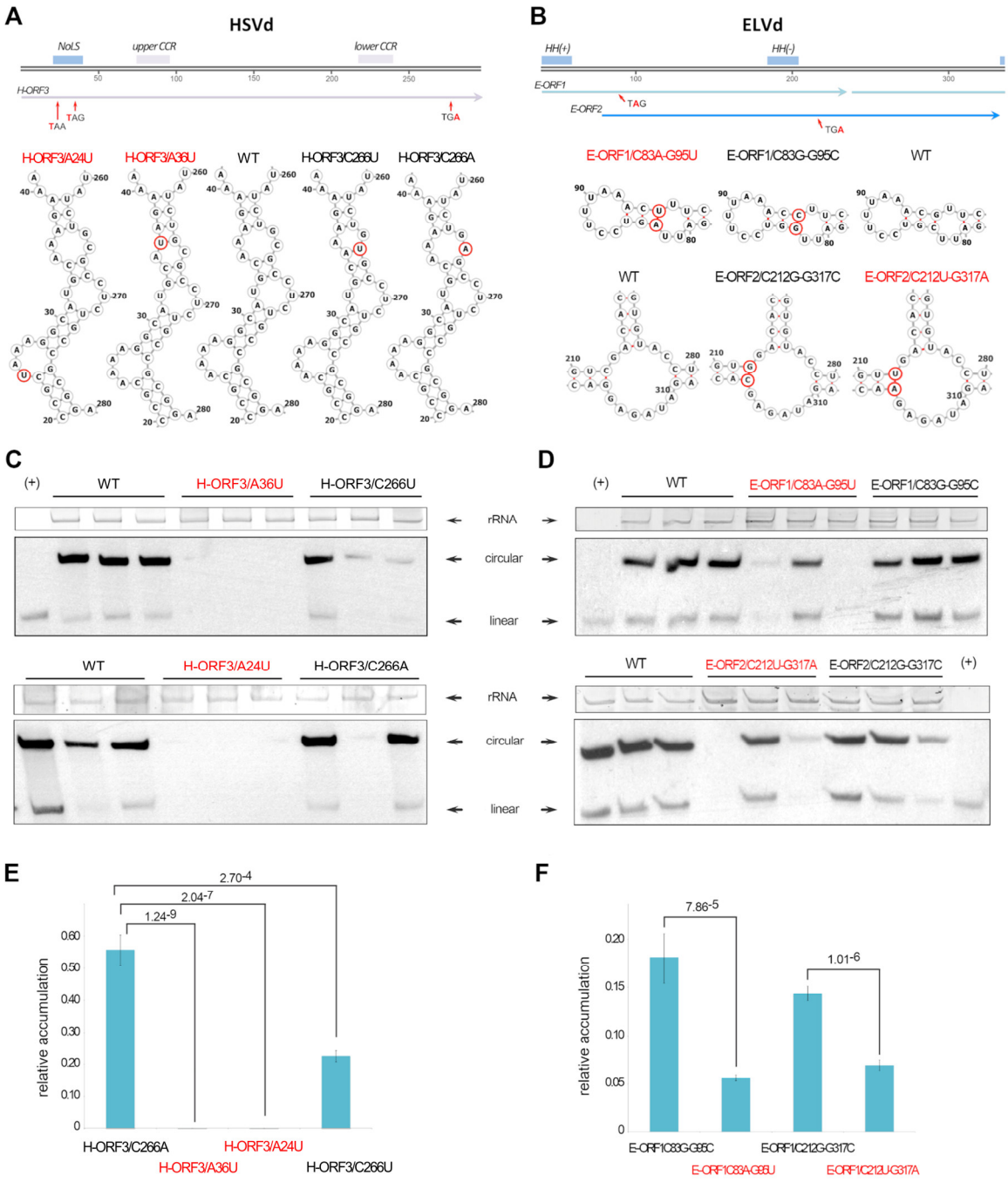


Figure 2.3. Truncated ORFs affect viroid biological efficiency. Representation of the HSVd (**A** - upper panel) and ELVd (**B** - upper panel) linear monomers (plus polarity). The Central Conserved Region (CCR) and the potential Nucleolar Localization Signal (NoLS) (in HSVd) and Hammerhead Ribozyme (HH) (in ELVd) are highlighted (boxes). The position of the nucleotides changed to introduce the stop codons in the conserved ORFs are marked with arrows. Predicted secondary structure of the regions of the HSVd (**A** - lower panel) and ELVd (**B** - lower panel) in which were performed the nucleotide substitutions (marked with circles and arrows). WT represents the structure of the unmodified HSVd and ELVd sequences. Representative northern blot hybridization of RNA extracted from cucumber an eggplant inoculated with the HSVd (**C**) and ELVd (**D**) mutant-variants. Ribosomal RNA (rRNA) was used as load control. HSVd and ELVd linear transcripts (+) were used as hybridization controls. Relative accumulation of HSVd (**E**) and ELVd (**F**) mutant-variants in inoculated plants, estimated by qRT-PCR. The statistical significance was estimated by paired T-tests and the obtained p-values are shown. Error bars represent the standard error values.

Dimeric constructs of HSVd and ELVd variants were used to inoculate cucumber and eggplant plants, respectively. At 21 dpi total RNAs were extracted to estimate viroid accumulation. Mature forms of H-ORF3/A24U and H-ORF3/A36U mutants (carrying stop codons in the R/K motif) were not detected by northern blot assays (Figure 2.3C – upper panel). In contrast, circular forms of H-ORF3/C266A and control mutants (although with variable accumulation levels) were identified in all the analyzed samples (Figure 2.3C – lower panel). These results were reinforced by the observation that the symptom intensity in H-ORF3/A24U and H-ORF3/A36U inoculated plants was significantly lower than that observed in cucumber plants inoculated with the unmodified and/or control (H-ORF3/C266U) variants (Supplementary Figure 2.9). Regarding ELVd, circular forms of all the analyzed viroid-sequences were detected by northern blot (Figure 2.3D). However, it was evident that the accumulation of both E-ORF1/C83A-G95U and E-ORF2/C212U-G317A variants with truncated ORFs was decreased in relation to control mutants (Figure 2.3D). The sequence stability of the mutant viroid variants in the analyzed plants was confirmed by sequencing 10 clones of each (Supplementary Figure 2.10), observing that only in the case of the C83A-G95U mutant one out of the ten clones reverted to the wildtype sequence revealing a high stability of the introduced mutations.

To perform a more precise determination of HSVd and ELVd levels, total RNAs from inoculated plants were also analyzed by qRT-PCR (Figure 2.3E). As expected, all HSVd mutants presented lower accumulation levels than the wild type, being the control sequence H-ORF3/C266U the less affected variant. In contrast, mutants H-ORF3/A24U and H-ORF3/A36U showed accumulation values near zero (Figure 2.3E). The HSVd mutant H-ORF3/C266A, carrying a stop codon out of the NoLS presented accumulation values intermediate between both NoLS-disrupted mutants and H-ORF3/C266U control. Results obtained from the HSVd qRT-PCR were similar to those of northern blot analysis (Figure 2.3F). These results revealed that the estimated levels of ELVd variants carrying stop codons in their sequence are significantly lower than the observed for their respective controls with unaltered ORFs.

Circular forms of HSVd and ELVd interact with translational machinery *in vivo* and HSVd contains m6A modification

It is generally assumed that the functional status of a translatable RNA is associated with the specific presence of such RNA in polysomes fractions. Consequently, we performed polysome profiling of viroid-infected plants to study the possibility that *ex*-circRNAs might be incorporated into plant translational complexes. Ribosome fractions from cucumber and eggplant leaves infected with HSVd and ELVd, respectively, were separated by ultracentrifugation on a sucrose density gradient, analyzed by absorbance at 260 nm and collected for posterior analysis (Figure 2.4A and B). In order to monitor the presence of *ex*-circRNAs, total RNA was extracted from cytosolic (2 and 3), and polysome-associated (10 and 11) fractions, which were discriminated by the position of the 80S peak that corresponds to monosomes. As positive control we analyzed endogenous mRNAs constitutively expressed in both analyzed plants. The negative control consisted of a linear transcript of another *ex*-circRNA (PSTVd) added to the extracts before being loaded into the sucrose gradient. RT-PCR analysis revealed the presence of internal endogenous controls and HSVd and ELVd RNA in all the analyzed fractions, whereas linear PSTVd transcripts were only detected in cytosolic fractions (Figure 2.4C and D). To confirm that circular forms of HSVd and ELVd were associated with polysome fractions, total RNAs extracted from fractions 10 and 11 were analyzed by northern blot assays. Circular forms of plus-strand HSVd and ELVd (marked with arrows in the Figure 2.4E and F) were clearly detected in the polysome fractions in both analyzed plants. These results confirm that both *ex*-circRNAs analyzed here, are associated *in vivo* with plant polysomes. Regarding to the detection of linear forms of HSVd and ELVd, at this point, we are incapable to determine if these detected transcripts are derived from circular RNAs broken during the purification process or constitute evidence that some linear monomers of both analyzed viroids might also be associated with polysomes.

One issue to be solved is how the translation of these viroid-derived peptides might start in the absence of an initial methionine. Nevertheless, there are a growing number of reports on alternative mechanisms of translation initiation (Kwan and Thompson, 2019). Among them, N6-Methyladenosine (m6A) in RRACH sites (being R [A or G] and H [A or C or U]) can promote internal initiation of polysome-associated endogenous circular RNAs (circRNA) in cells (Yang et al., 2017). In this regard, canonical RRACH sites were computationally identified in HSVd and ELVd (Supplementary Figure 2.11, upper and lower panel respectively). Moreover, immunoprecipitation of total RNA extracted from HSVd-infected cucumber plants revealed the presence of the plus polarity of HSVd in the RNAs immunoprecipitated with m6A antibody but not in those RNAs recovered after immunoprecipitation with GFP antibody, which was used as negative control (Supplementary Figure 2.11B). Besides, circular forms of HSVd were isolated,

digested into nucleosides and analyzed by UPLC-PDA-Tof-MS. A putative m6A peak (recovered at 6,48 s) was identified in comparison with the pattern obtained with a commercial N6-methyladenosine marker (Supplementary Figure 2.11C).

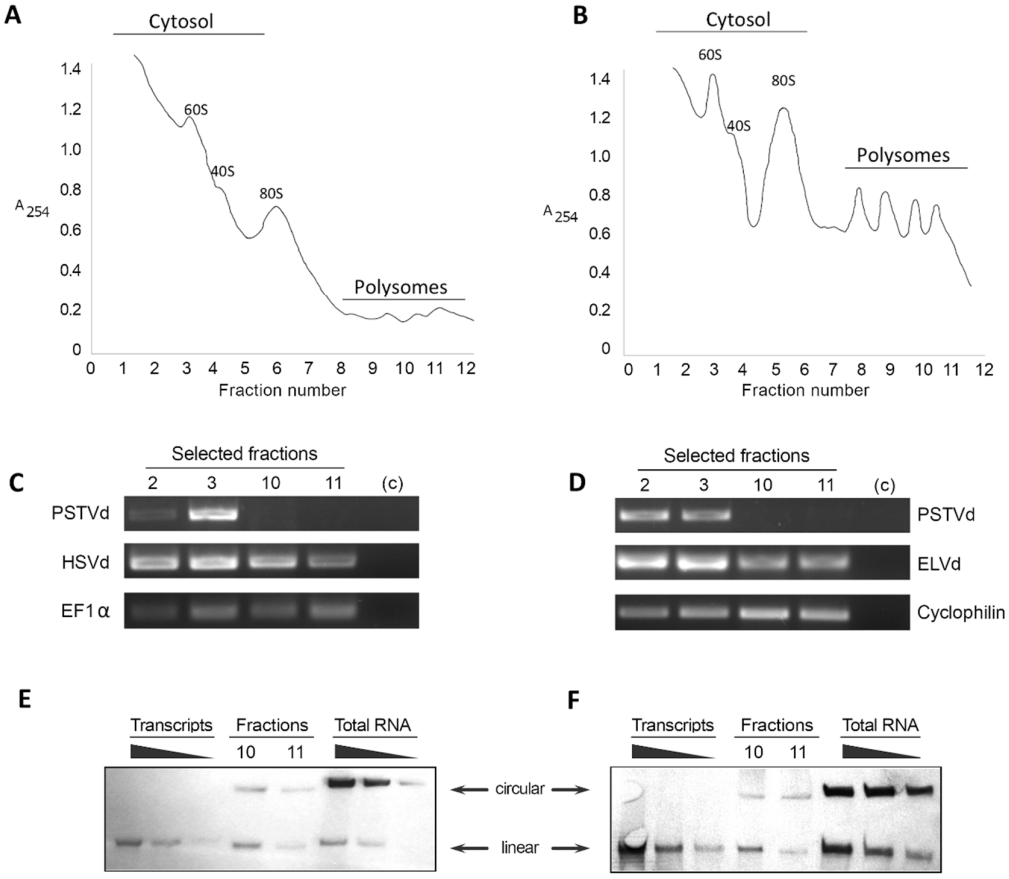


Figure 2.4. *Ex*-circRNAs are associated with translational machinery. Sucrose density gradient analysis of fractions from leaves infected with HSVd (**A**) or ELVd (**B**). The position of the 80S monosomes is indicated to separate the polysome fractions from the cytosolic ones. Total RNAs extracted from fractions 2 and 3 (cytosol), and 10 and 11 (polysomes) of HSV- and ELVd-infected plants (**C** and **D**, respectively) were analyzed by RT-PCR. Detection of EF1α (in cucumber) and Cyclophilin (eggplant) transcripts was used as endogenous control. An *in vitro* generated transcript of PSTVd (added to the plant extract before being loaded into the sucrose gradient) was used as negative control. HSVd and ELVd RNA and endogenous controls were detected in all analyzed fractions. In contrast, exogenous control (PSTVd) expected to be not associated with polysomes, was only present in cytosolic fractions. Northern blot of RNA isolated from polysome fractions derived from leaves infected with HSVd (**E**) or ELVd (**F**). Serial dilutions of monomeric transcripts and total RNAs extracted from infected plants were used as hybridization controls. Analyses were performed in duplicate and only representative results are shown.

Discussion

CircRNAs were originally regarded as a novel class of ncRNAs. However, recent studies have provided evidence that circRNAs can also act as non-canonical translatable transcripts (Diallo et al., 2019). Different functional peptides derived from circRNAs have been described in animal cells (Legnini et al., 2017; Pamudurti et al., 2017; Yang et al., 2017; Zhang et al., 2018b; Zhang et al., 2018c; Liang et al., 2019; Zheng et al., 2019a), supporting the emerging notion that the true coding-potential of circRNAs may have been underestimated. Although circRNAs were described initially in plants (Sanger et al., 1976), research on this topic is still in its infancy as no autonomous coding plant-circRNA has been described yet (Zhao et al., 2019). Given this new scenario, we used HSVd and ELVd, two well-known pathogenic circRNAs representatives of the two viroid families characterized by replicating and accumulating in nucleus and chloroplasts, respectively, as experimental tools to explore the coding potential of circRNAs in plants.

Computational analysis revealed the existence of conserved ORFs in both analyzed *ex*-circRNAs. H-ORF3 and E-ORF1 were the predicted peptides showing the highest conservation rates in HSVd and ELVd accessions, respectively. It is important to consider that ELVd is a more recently described viroid and that the reported variants come from a unique host (eggplant). In contrast, HSVd is a viroid largely studied with multiple sequence-variants recovered from diverse host-plants and (in some cases) submitted to sequence repositories without exhaustive previous quality controls (for example, 73 sequences were excluded from this analysis because they contained some nucleotide indeterminacy and/or incomplete sequence). Consequently, it cannot be excluded the possibility that the conservation rate inferred for H-ORF3 (lower than that observed for E-ORF1 and E-ORF2) can be underestimated.

The study of specific cellular localization has been frequently assumed as key knowledge for elucidating biological activity and protein functions in the cell (Nair and Rost, 2003; Lin et al., 2009). Consequently, the observation that peptides encoded by ORFs detected in *ex*-circRNAs analyzed here show predominant localization in the nucleolus (H-ORF3) and chloroplasts (E-ORF1) provides the primary observation suggesting a potential functional role. Interestingly, the specific subcellular compartmentalization of these putative peptides is coincident with the organelles where both plant-pathogenic RNAs predominantly accumulate, nucleolus for a member of the family *Pospiviroidae* (HSVd) and chloroplasts for a member of the family *Avsunviroidae* (ELVd). Previously, it has been suggested that the specific compartmentalization of ELVd-derived constructs may be regulated by the viroid-RNA sequence (Gómez and Pallás, 2010a). However, the coding potential of those sequences could have been undervalued because of the primary consideration as plant-pathogenic ncRNAs. Remarkably, ELVd contains an additional ORF (E-

ORF2), potentially encoding a peptide with redundant biological activity (localization in chloroplasts). It has been demonstrated that viroids in the family *Avsunviroidae* exhibit the highest mutation rate reported for any biological entity (Gago et al., 2009), whereas the mutation rates of pospiviroids are significantly lower (López-Carrasco et al., 2017). Under this strict evolutionary constriction, it may be expected that, according to the role of genetic redundancy in robustness (Fares, 2015), the functionality played by a determined component (E-ORF1 encoded peptide, in this case) may be guaranteed by another (E-ORF2) with a functional overlap. Finally, the additional identification of ORFs potentially encoding peptides carrying functional NoLS in a representative member of the family *Pospiviroidae* (PSTVd), provides additional support to the notion that potential circRNA-derived peptides inferred here, might possess biological activity.

The protein-coding capability of a small sub-set of endogenous circRNAs in animals has been validated by the identification of circRNA-encoded peptides by mass spectrometry (MS) (Patop et al., 2019). Unfortunately, our attempts to detect the predicted peptides by Liquid Chromatography MS (LC-MS) in total protein-extracts recovered from infected tissues were unsuccessful. However, it could not be excluded the possibility that the analyzed tissues could represent a mix of infected and non-infected cells (Harders et al., 1989), thus prompting that peptide accumulation might be below the LC-MS detection limits in crude extracts (Karpievitch et al., 2010). Therefore, additional peptide-enrichment steps (Liang et al., 2019) and/or single-cell proteomic analysis (Shi et al., 2018) could be needed in order to unequivocally confirm, in the future, the presence in infected-plants of the putative circRNA-derived peptides described here. Searching for biological evidence supporting the involvement of circRNA-derived peptides in the infection cycle of viroids, we constructed diverse mutant-variants that truncate the identified ORFs. Infection bioassays revealed that variants carrying a stop codon in their sequence exhibit a significant lower accumulation in comparison with control mutants. The most evident effect was observed in cucumber plants inoculated with HSVd, where mature forms of H-ORF3/A24U and H-ORF3/A36U were practically undetectable. Regarding ELVd mutants, our results revealed that although ORF-truncated variants showed lower accumulation levels than their corresponding controls, the biological effect of the introduction of stop codons in both mutants was less evident than the observed in HSVd. At this point, it is important to remark that according to our prediction, ELVd possesses two putative ORFs with hypothetical functional activity, consequently in any of the two analyzed mutants we were not able to generate a true potentially non-functional variant.

Diverse studies have shown that *translatome* technologies are the most reliable method to find coding circRNAs (Legnini et al., 2017; Pamudurti et al., 2017). The demonstration that circular forms of HSVd and ELVd are physically associated with translating polysomes, reveals their

involvement in the plant translational machinery as recently described for the circular forms of the citrus exocortis viroid (another member in the family *Pospiviroidae*) (Cottilli et al., 2019). In this latter work authors suggested that the interaction viroid-ribosome is related to the pathogenesis process in infected tomato plants as a consequence of ribosomal stress. However, the demonstration that besides HSVd (a viroid causing strong phenotypic alterations), circular forms of the ELVd also interact with the polysomes in non-symptomatic infected eggplant, suggests that this *in vivo* association between *ex*-circRNAs and the plant translational machinery may possess additional biological functions beyond the physiological alteration manifested as symptoms expression in infected plants.

The coding potential of circRNAs raises the interesting question of which mechanisms regulate the translation initiation of those covalently closed transcripts lacking 5′cap structure (Diallo et al., 2019). Although the predominant mode of translation initiation is cap-dependent, translation does not always start at the first AUG encountered by the ribosome and diverse alternative initiation mechanisms (re-initiation, initiation at non-AUG codons, internal ribosomal entry, N^6-Methyladenosine (m6A), etc.) have also been described in plants (Merchante et al., 2017) and/or mammalians (Hou et al., 2020). There are also cases in which the mechanisms regulating translation in plants remain unknown (Hansen et al., 2001; Hua et al., 2001). Interestingly, these non-canonical translation initiation processes are generally activated under stress (Diallo et al., 2019), like that induced in the plant cell as a consequence of viroid infection. The prediction of canonical RRACH domains (the sequence context in which m6A modifications predominantly occur in mRNAs) (Yang et al., 2018; Arribas-Hernández and Brodersen, 2020) in HSVd and ELVd RNAs (Supplementary Figure 2.11A), and the detection of potential m6A residues in HSVd circular RNAs (Supplementary Figure 2.11B and C), allow us to speculate about the existence of non-canonical initiation mechanisms of translation in viroid RNAs, perhaps dependent on m6A modification as it has been reported for circRNAs of mammals (Yang et al., 2017; Diallo et al., 2019; Prats et al., 2020). However, further studies are needed to decipher the mechanistic aspects of how the potential translation of *ex*-circRNAs would be regulated in plants.

In our opinion, these results generated through computational approaches, assays of subcellular compartmentalization, analysis of functional activity and polysome profiling, provide novel insights about the possible translation of exogenous plant-circRNAs (using viroids as experimental models) and contributed to expand the coding landscape of the cell transcriptome, suggesting the existence of an unexplored layer of gene activity in plants related to the potential protein-coding circRNAs, whose biological functions are largely yet to be revealed (Zhao et al., 2019).

We are aware that the non-coding nature of circRNAs in general and viroids in particular constitutes a dogma that is difficult to refute, but progress showing that some animal circRNAs can code for small peptides (Diallo et al., 2019; Patop et al., 2019) and the recent demonstration that CEVd causes ribosomal stress in tomato plants (Cottilli et al., 2019) are emerging insights that prompted us to reconsider this notion. Here, we provide experimental observations favoring that the coding capacity of viroids cannot be ruled out, although the definitive evidence (detection of the circRNA-encoded peptides) is a technological challenge to be addressed in future research lines focused on deciphering the molecular biology of the circular RNAs in plants.

Material and methods

In silico peptide identification

Sequences of the positive strand of ELVd (AJ536613.1) and HSVd (Y09352.1) were translated into the three possible reading frames using Virtual Ribosome (Wernersson, 2006). The parameters used were standard genetic code, all codons were considered as internal and stop codons were read through to obtain several polypeptides of the same reading frame. Generally, peptides derived from small ORFs in plants have an average length ranging between 50 to 100 amino acids (Couso and Patraquim, 2017). Consequently, we focused our study on the putative viroid-derived peptides with a length > 45 amino acids, and therefore similar to what has been previously described in plants.

Genomic sequences of 104 and 779 accessions, respectively, for ELVd and HSVd, were downloaded from NCBI. Each sequence was duplicated to account for the circular nature of the viroid genome. The positive strand of the obtained dimeric sequence was translated into the three possible reading frames using Virtual Ribosome as described previously (Wernersson, 2006). Those peptides were aligned using ClustalΩ default parameters against the ORFs identified in the infectious clones of ELVd (AJ536613.1) and HSVd (Y09352.1). The identity values were inferred from the similarity scores computed in the identity matrix, which was automatically rendered by ClustalΩ. Only the identity scores of all peptides compared to the reference ORFs were used to calculate the mean identity values. The conservation of the putative R/K motif (RKEKRRR) present in PSTVd RG1 (U2305.1) was inferred using complete genomic sequences of PSTVd (402) from NCBI, which were analogously translated using Virtual Ribosome. A pairwise alignment of the R/K motif to the peptide sequences was implemented using the local alignment function hosted in the "pairwise2" Biopython module, allowing for up to two mismatches. All Logo representations were obtained using WebLogo (Crooks et al., 2004).

Construction of fusions to fluorescent proteins

Peptides derived from the three possible reading frames of ELVd (AJ536613.1) (Fadda et al., 2003) and HSVd (Y09352.1) (Kofalvi et al., 1997) were cloned upstream of a fluorescent protein into a pBluescript SK vector digested with *Nco*I and *Nhe*I. Specifically, ELVd-derived sequences were cloned in frame with yellow fluorescent protein (YFP), whereas HSVd-derived sequences were cloned to GFP. DNA fragments were amplified with high-fidelity PrimeSTAR HS DNA Polymerase (Takara, Kusatsu, Japan) and overhang ends were generated by digestion with the IIS restriction enzyme *Bsa*I. Moreover, the Fibrillarin gene of *N. benthamiana* (AM269909.1) was cloned upstream mCherry in an analogous manner. Nucleolar signals were cloned upstream GFP using self-complementary primers that upon hybridization generate overhang ends compatible with

*Nco*I in the 5' and with *Nhe*I in the 3'. The pBluescript SK vectors used for the subcloning in frame with the different fluorescent proteins contain a duplicated cauliflower mosaic virus 35S promoter, a 5' UTR of alfalfa mosaic virus and a Popit terminator (Supplementary Figure 2.2). The resulting transcriptional units were cloned into the binary vector pMOG 800 using *Hind*III and transformed into agrobacterium C58. All oligonucleotides used for generating these constructs are listed in Supplementary Table 2.1.

Confocal microscopy

Nicotiana benthamiana leaves of 3-4 weeks old plants were agroinfiltrated with a culture of *A. tumefaciens* strain C58, transformed with the performed construct. The overnight grown bacterial culture was diluted in infiltration buffer (MES 0.1 M, $MgCl_2$ 1 M) up to an optical density at 600 nm (OD600) of 0.2 and injected on the abaxial side of the leaves using a 1 ml needle-less syringe. Agroinfiltrated plants were grown at 25 °C (light) and 18 °C (darkness) and analyzed two days post-agroinfiltration.

Subcellular localization of the fluorescence and imaging was conducted with an inverted Zeiss LSM 780 confocal microscope and ZEN software (Carl Zeiss). Images were acquired using an objective plan-apochromat 40x/1.4 Oil DIC M27. 0.5 cm diameter leaf dishes were cut and mounted in water. The chlorophyll and the YFP were excited with the 488 nm and 514 nm lines of an argon ion multilaser, respectively. The YFP emitted fluorescence was collected between 530 and 570 nm and chlorophyll autofluorescence was collected from 700 nm and beyond. We ensured that there was no cross-talk between the YFP and chlorophyll signals by using sequential instead of simultaneous scanning settings. GFP was excited with the 488 nm laser and mCherry with 594 nm laser.

Western Blot

Proteins were extracted from agro-infiltrated *N. benthamina* leaves homogenized with Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 15% glycerol, 2% SDS, 350 mM DTT). Samples were boiled for 10 minutes and centrifuged at 13000g for 15 minutes. The supernatant was loaded into a 12% acrylamide gel and the SDS-PAGE was run for two hours using the MiniProtean 3 system (BioRad). Subsequently, it was transferred to a polyvinylidene fluoride (PVDF) membrane (GE Healthcare Life Sciences). Immunodetection was performed using a primary mouse antibody against the GFP 1:10,000, and an anti-mouse secondary antibody conjugated with peroxidase at 1:10,000. For chemiluminescent detection Pierce™ ECL Western Blotting Substrate was employed and LAS3000 for image acquisition.

Construction of mutant viroid clones

Point mutations were designed according to the following viroids-sequence limitations: *i*) sequence changes must be only substitutions (in order to maintain the sequence length), *ii*) nucleotides should be located in a sequence context allowing the generation of stop codons with a single base substitution, *iii*) substitutions must not involve nucleotides located in highly conserved domains (i.e. Conserved Central Region (CCR) and Terminal Conserved Hairpin (TCR) for HSVd and plus and minus Hammerhead ribozymes (HH) for ELVd) considered strictly necessary for viroid life cycle, and *iv*) nucleotide modifications (in both truncated ORFs and control mutants) must be conservative with the predicted secondary structure of the viroids. RNA secondary structures files were generated in Vienna format and were displayed using Forna (Kerpedjiev et al., 2015). HSVd (Y09352.1) secondary structure of the plus polarity was predicted using default parameters of Mfold (Zuker, 2003) and RNAfold (Gruber et al., 2008). The resultant structure was the characteristic rod-like conformation of the members of *Pospiviroidae* family. Point mutations were designed in the part of the structure that is coincident between both *in silico* predictions (positions 1-70 and 228-297).

The point mutations (Figure 2.2) were introduced using mutagenic primers by amplifying the entire monomeric viroid sequence with high-fidelity PrimeSTAR HS DNA Polymerase (Takara, Kusatsu, Japan). In particular, H-ORF3/A24U and H-ORF3/A36U mutations were generated by PCR with a forward primer carrying the mutation (Fw H-ORF3/A24U and Fw H-ORF3/A36U, respectively) and Rv D2-HSVd, while H-ORF3/C266A and H-ORF3/C266U with Fw D1-HSVd and a reverse primer with the mutation (Rv H-ORF3/C266A and Rv H-ORF3/C266U, respectively). The mutated monomers were dimerized and cloned into the binary vector pMD201t as previously described (Marquez-Molins et al., 2019). All oligonucleotides used for generating these constructs are listed in Supplemental Table 2.1.

ELVd (AJ536613.1) secondary structure of the plus polarity had been previously elucidated and the *in vivo* data obtained by high-throughput selective 2'-hydroxyl acylation analyzed by primer extension (hSHAPE) support the prediction of Mfold software (Zuker, 2003). The point mutations conservative with secondary structure were introduced by site-directed mutagenesis. Dimeric ELVd cloned into a pBluescript SK vector was amplified with self-complementary mutagenic primers, treated with *Dpn*I and transformed into competent *E.coli* cells (Sanjuán and Daròs, 2007). Dimeric forms were selected by colony PCR using primers of the plasmid (Fw 35S and Rv Popit). Another step of site-directed mutagenesis was required for ELVd-C212G-G317C and ELVd-C212U-G317A mutants. Oligonucleotides used for generating these constructs are listed in Supplemental Table 2.1. An *in vitro* transcription assay was performed to validate that the hammerhead activity of ELVd-mutants has not been altered in comparison with the wild type

(AJ536613.1). For this purpose, dimeric sequences were amplified to introduce a T7 promoter just before the beginning of the viroid sequence, using primers (Fw T7 965 and Rv 966). Transcription was performed with T7 RNA polymerase (Takara, Kusatsu, Japan) for 5 minutes and stopped by adding 1x volume of inhibition solution (50 mM EDTA, 8 M urea and 50% (w/v) formamide) and then loaded into a 1% agarose gel.

Plant material and viroid inoculation

Cotyledons of *Cucumis sativus* cv. Marketer and *Solanum melongena* cv. Black Beauty were agro-infiltrated with a culture of *A. tumefaciens* strain C58 harboring the correspondent construct. The overnight grown bacterial culture was diluted in infiltration buffer (MES 0.1 M, $MgCl_2$ 1 M) up to an optical density at 600 nm of 1 and injected on the abaxial side of both cotyledons using a needle-less syringe. Plants were kept in a photoperiod of 16 hours under visible light and 30 °C (light)/25 °C (darkness) for *C. sativus* and 25 °C (light) /18 °C (darkness) for *S. melongena*. Samples of systemic leaf tissue were collected at 21 days after viroid inoculation. Specifically, half of the second and third leaves after the apex.

Northern blot

Total RNA was extracted from systemic leaves using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). For Northern Blot analysis 3 µg of total RNA per sample were mixed with solid urea and then loaded into a PAGE 5% urea 8 M and TBE 89 mM gel. RNA electrophoresis was performed at 200 V for 1 h and then RNA was transferred to a nylon membrane using MiniProtean 3 system (BioRad, Hercules, CA, USA). Transfer conditions were 100 V for 1 h at 4 °C in TBE buffer 1×. Nucleic acids transferred to the membrane were covalently fixed by using ultraviolet light (700×100 J/cm^2). Hybridization and chemiluminescent detection were performed as previously described (Herranz et al., 2005).

RT-qPCR

First-strand cDNA was synthesized by reverse transcription using random hexamers and RevertAid cDNA Synthesis Kit (Thermo Scientific™, Waltham, MA, USA). qRT-PCR assays were performed using PyroTaq EvaGreen mix Plus (ROX) (CulteK Molecular Bioline) according to the manufacturer's instructions. All analyses were performed in triplicate on an ABI 7500 Fast-Real Time qPCR instrument (Applied Biosystems) as previously described (Sanz-Carbonell et al., 2020). The efficiency of PCR amplification was derived from a standard curve generated by four 10-fold serial dilution points of cDNA obtained from a mix of all the samples. RNA expression was quantified by the comparative ΔΔCt method. Primers for amplifying viroid sequences were

picked in regions that excluded the mutagenized positions and were blasted against the host genome to guarantee that do not amplify any plant gene. The sequence of all primers is in Supplemental Table 2.1. Relative RNA expression was determined by using the comparative ΔΔCT method (Livak and Schmittgen, 2001) and normalized to the geometric mean of ubiquitin (NM-001282241.1) for cucumber and cyclophilin (JX448344) for eggplant expression, as reference. The statistical significance of the observed differences was evaluated by the paired t-Test.

Polysome profiling

Polysomes were isolated from infected plants by ultracentrifugation on a sucrose density gradient as previously described (Lecampion et al., 2016). Briefly, 200 mg of leaf tissue (per sample) were homogenized in 1.6 ml of precooled polysome extraction buffer (160 mM Tris-HCl pH 8.4, 80 mM KCl and 40 mM $MgCl_2$, 5.26 mM EGTA, 0.5% (w/v) Triton X-100, 50 $\mu g.ml^{-1}$ cycloheximide, 50 $\mu g.ml^{-1}$ chloramphenicol) and centrifuged at 16,000 x g for 15 min at 4 °C. The supernatant was filtered with miracloth (Sigma-Aldrich, St Louis, MI, USA), 1860 pg of PSTVd dimeric transcript were added as internal control, homogenized, and then loaded on top of a sucrose gradient (20-50% (w/v)). Ultracentrifugation conditions were 175,000 x g for 2 hr 45 min at 4 °C. Subsequently, 1 mL fractions were collected and the absorbance at 254 nm was measured. RNA was extracted from sucrose fractions adding 1 volume of phenol/chloroform/isoamyl alcohol (25:24:1), mixing thoroughly and centrifuging at 14,000 x g for 10 min at 4 °C. The aqueous phase was recovered, and this step was repeated with chloroform/isoamyl alcohol (24:1). RNA was precipitated with absolute ethanol and resuspended in 20 µL of RNase-free water. One-step RT-PCR was performed with Superscript™ III and Platinum® Taq (Thermo Scientific™, Waltham, MA, USA). Reactions were carried out with 150 ng of RNA, 5 µl 2x Reaction Mix 0.5 µl of each primer (10 µM), 0.4 µl of SuperScript™ III RT/ Platinum ® Taq enzyme mix, 0.1 µl RiboLock RNase inhibitor (40 U/µL) and nuclease-free water up to 10 µl. The primers used are listed in the Supplemental Table 2.1. RT-PCR conditions were an initial step of 30 min at 45 °C followed by 25 PCR cycles of denaturation (94 °C), annealing (60 °C) and extension (72 °C). The duration of all steps in the cycles was 30 s. Northern blot analysis was performed as described above, except for the quantity of RNA (Figure 2.4).

m6A RNA-immunoprecipitation (RIP) of viroid genomic RNAs.

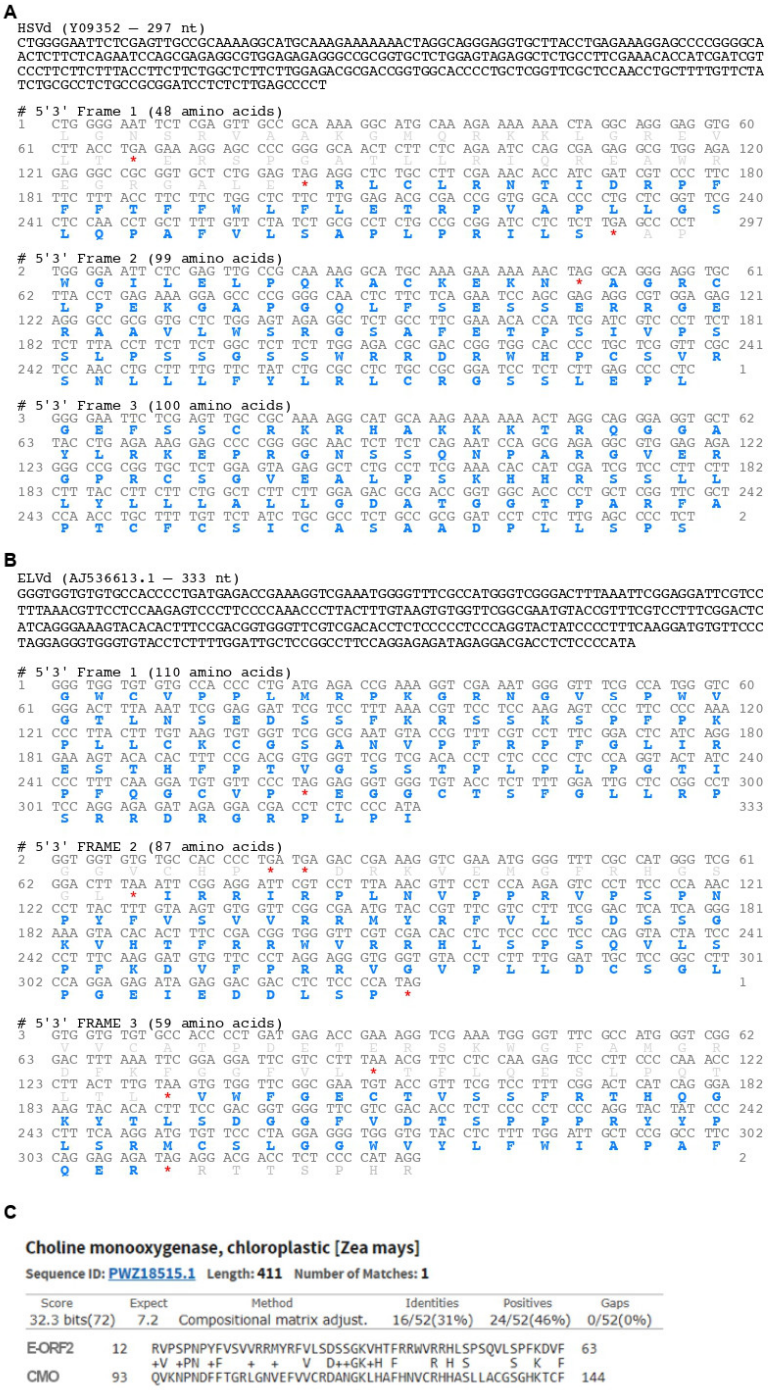
Immunoprecipitation of m6A viroid RNA was performed as previously described (Martínez-Pérez et al., 2017). Purified RNAs (200 µg) of *C. sativus* infected with HSVd were incubated in a final volume of 500 µL with 1× IP buffer [50 mM Tris·HCl, 150 mM NaCl, 0.5% (vol/vol) Nonidet P-40], 200 U RNasin and 12.5 µg of specific m6A antibody (Synaptic Systems) in rotation for 2 h at

4 °C. 0.1% of that volume was kept as an input. Then, 200 µL of protein A dynabeads in 1× IP buffer were added and additionally incubated in rotation for 2 h at 4 °C. After this, beads were washed four times with 1× IP buffer supplement with Ribolock and incubated for 1 hour with vigorous shaking in elution competition buffer [50 mM Tris·HCl, 150 mM NaCl, 0.5% (vol/vol) Nonidet P-40, 200 U Ribolock and 6.7 mM m6A] to elute m6A containing RNAs. As a negative control, the same reactions, using an unrelated antibody (IgG-anti GFP) instead of the m6A antibody, were performed. Eluted RNAs and serial dilutions (1/2, 1/10, 1/20) of them and the input were directly blotted onto a nylon membrane and hybridized with a specific probe to detect the plus polarity of the viroid.

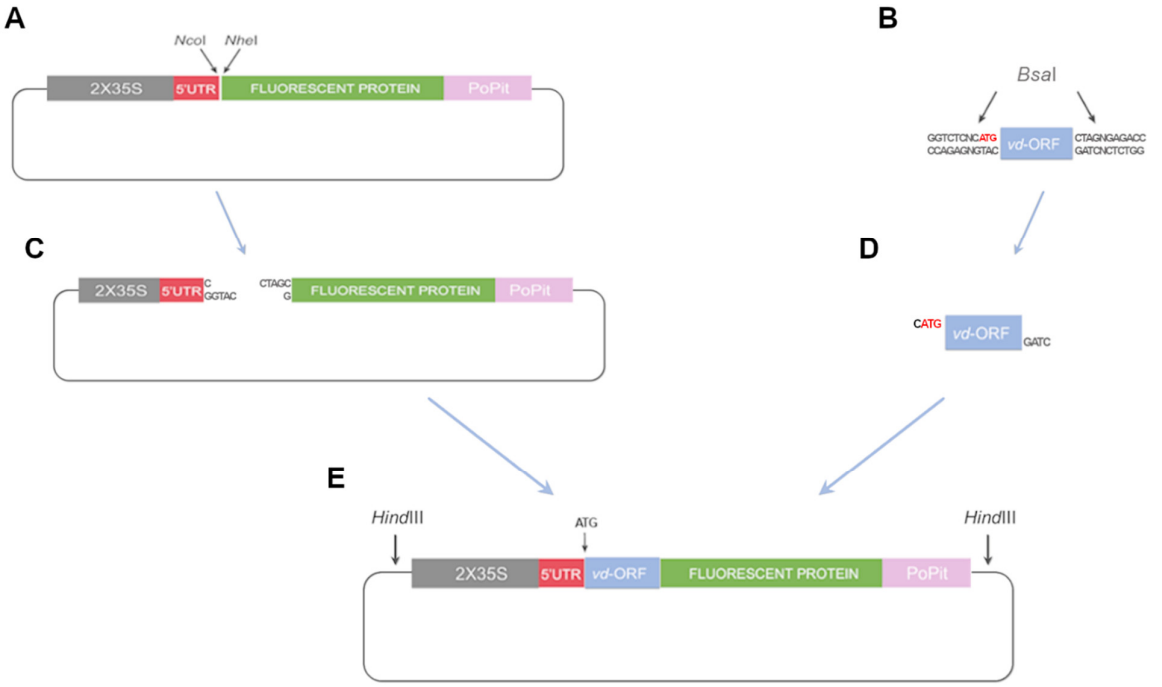
UPLC-PDA-Tof-MS Analysis.

Total nucleic acids were extracted with phenol and enriched in highly structured RNAs by a chromatography on CF11 cellulose as reported previously (Pallas et al., 1987). Subsequently, viroid circular RNAs were isolated by combining two polyacrylamide gels, the first under native conditions and the second denaturing. For UPLC-PDA-Tof-MS analysis, 2.5 µg of circular viroid RNA, extracted from the second polyacrylamide gel, was digested by incubating for 2 h at 37 °C in 50 µL (final volume) reaction mixture containing 1× buffer C (25 mM Tris pH 8, 2 mM MgCl2, 1 mg/mL; BSA), 10 U Benzonase, 0.002 U phosphodiesterase I and 1.5 U alkaline phosphatase. Then, 5 µL of digested nucleosides was injected into a Micromass Q-TOF spectrometer coupled to an Acquity UPLC-PDA system (Waters) via an electrospray ionization (ESI) interface. Separation was performed on a Waters Acquity BEH C18 column (150 × 2.1 mm i.d., 1.7 µm). During sample running, the mobile phase consisted of 0.1%formic acid in water (phase A), and 0.1% formic acid in acetonitrile (phase B). The solvent gradient program is conditioned as follows: 100–90% solvent A over the first 20 min, 90–0% solvent A over 10 min, return to the initial 100% A in 5 min, and conditioning at 100% A. The flow rate was 0.4 mL/min and the column and sample temperatures were kept at 40 °C. UV spectra were acquired between 210 and 800 nm with a 1.2-nm resolution and 20 points s−1 sampling rate. The ESI source was operated in positive ionization mode with the capillary and cone voltages at 2.7 kV and 30 V.

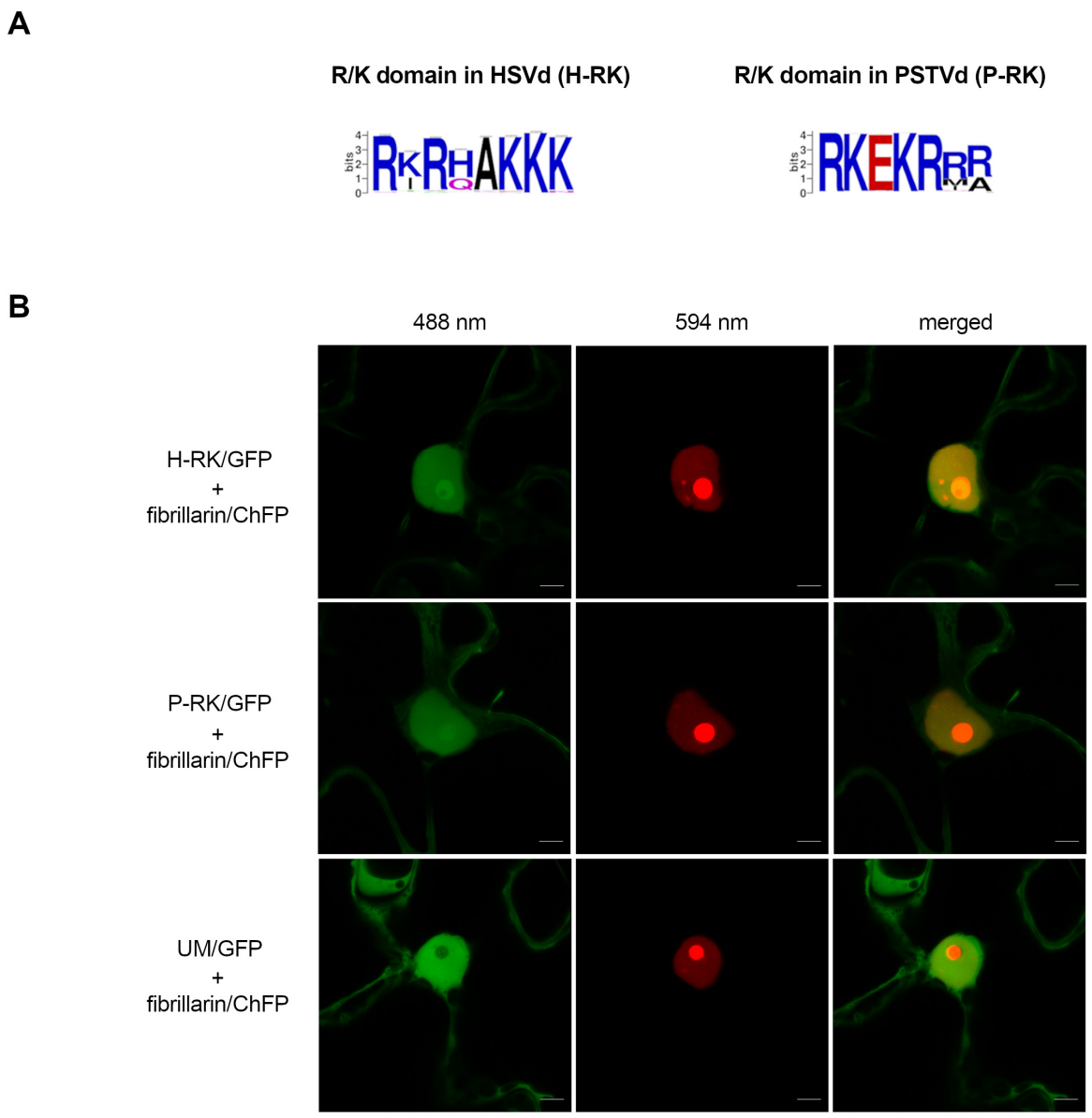
Supplementary material



Supplementary Figure 2.1: Detail of the ORFs predicted in ELVd and HSVd infectious clones. The positive strand of the HSVd (Y09352.1) (**A**) and ELVd (AJ536613.1) (**B**) accessions were translated into the three possible reading frames using Virtual Ribosome. Peptides were filtered by a minimum length of 45 amino acids and a conservation rate >50%. Selected ORFs are in blue. Stop codons are marked with red asterisks. **C**) C) Detail of the alignment of ORF-2 (performed with Blast-P -https://blast.ncbi.nlm.nih.gov- under default parameters) against higher plants database (taxid: 3193).



Supplementary Figure 2.2. Cloning strategy. A) pBluescript SK vectors carrying fluorescent protein cDNA were digested with NcoI and NheI. **B**) HSVd and ELVd sequences were amplified by PCR and digested with BsaI to generate compatible ends. Digested vector (**C**) and viroid-derived inserts (**D**) were ligated to generate constructs carrying viroid-derived peptides fused in frame with the fluorescent proteins (**E**). Sequence of the additional start codon is in red. Diagrams are not at scale.

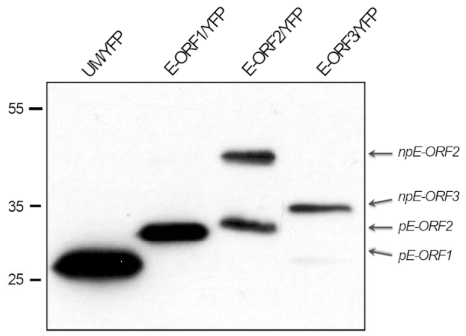


Supplementary Figure 2.3: Basic domains found in HSVd and PSTVd ORFs act as true NoLS. A) Detail of the putative R/K domains conserved in HSVd (H-RK) and PSTVd (P-RK) variants. The P-RK domain was found in 315 of the 402 PSTVd complete sequences available in public repository. **B**) Confocal microscopy images of *N. benthamiana* leaves agro-infiltrated with constructs expressing H-RK and P-RK domains fused to GFP. Unmodified GFP (UM/GFP) was used as a control. The fusion fibrillarin-ChFP was used as a nucleolus marker. Scale bars are 5 µm.

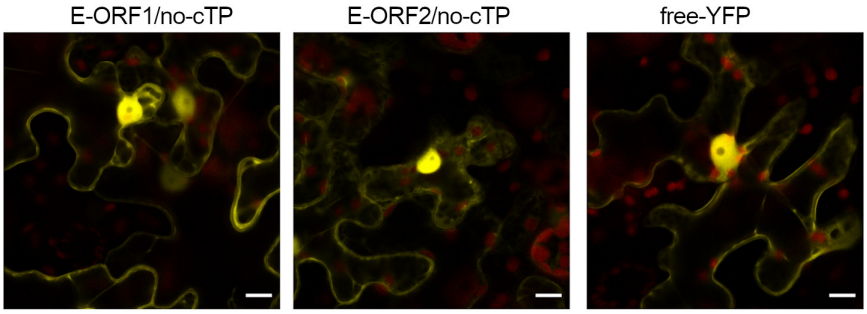
A

```
### chlorop v1.1 prediction results ##########################
Number of query sequences: 3

Name                    Length      Score  cTP      CS-score  cTP-length
-------------------------------------------------------------------------
E-ORF1                     110      0.506   Y         1.099          68
-------------------------------------------------------------------------

Name                    Length      Score  cTP      CS-score  cTP-length
-------------------------------------------------------------------------
E-ORF2                      87      0.550   Y        -1.189          73
-------------------------------------------------------------------------

Name                    Length      Score  cTP      CS-score  cTP-length
-------------------------------------------------------------------------
E-ORF3                      59      0.459   -         0.000          61
-------------------------------------------------------------------------
```

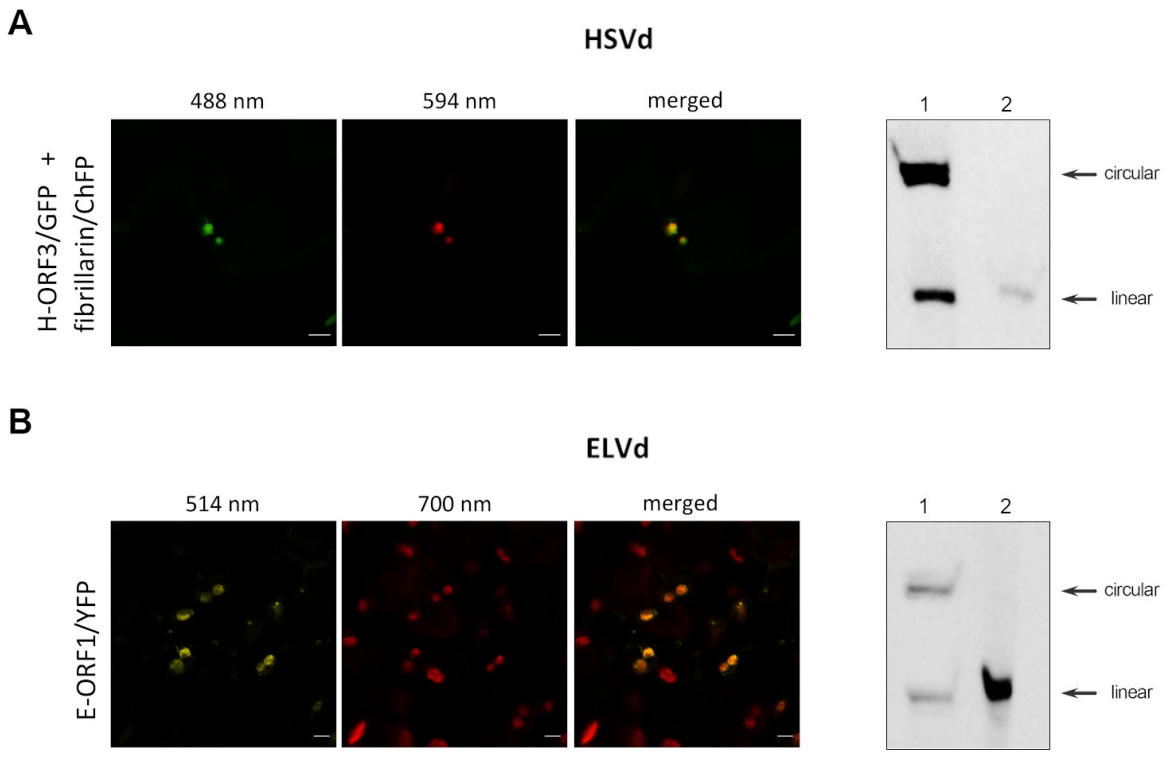
B



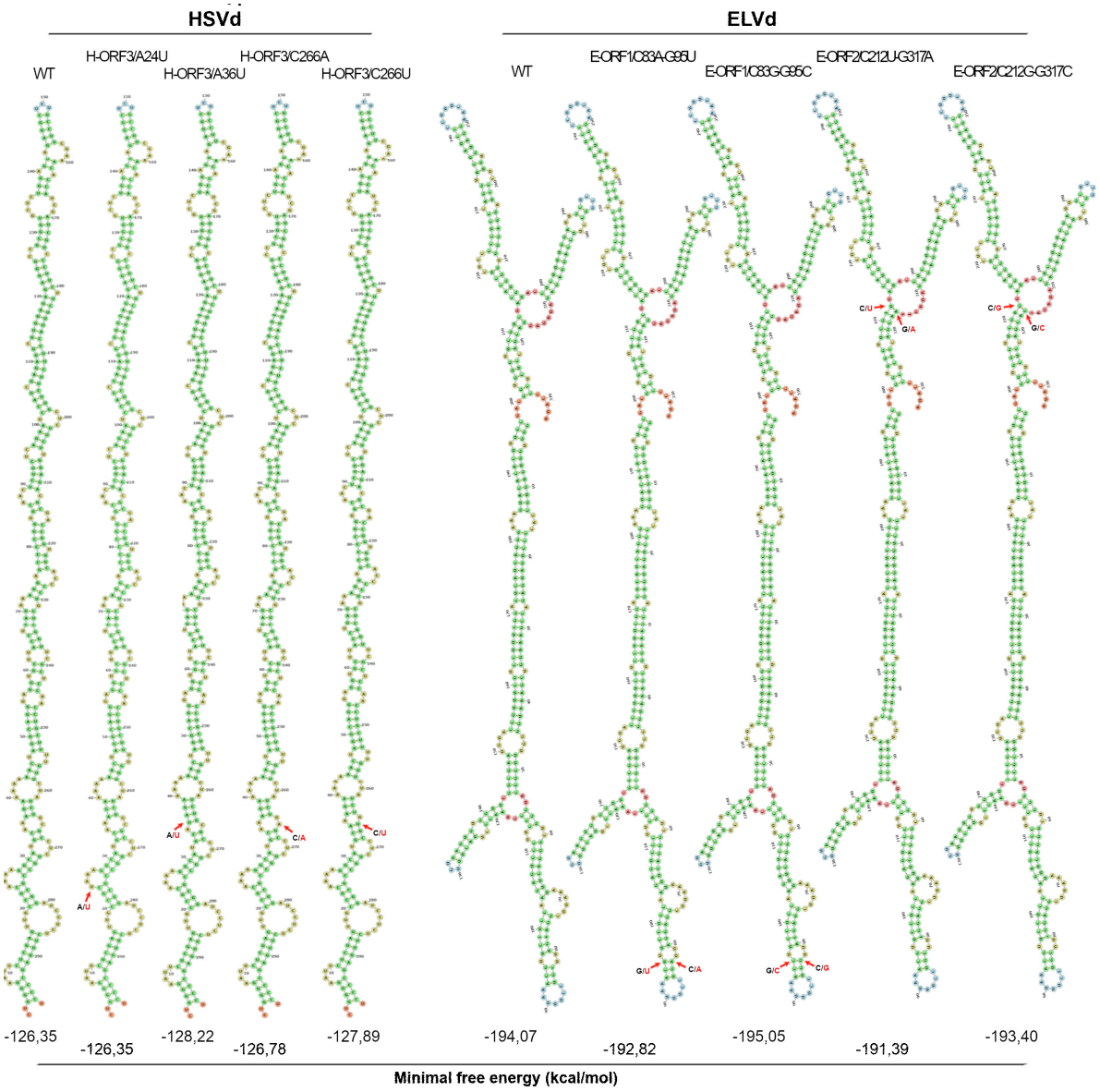
C



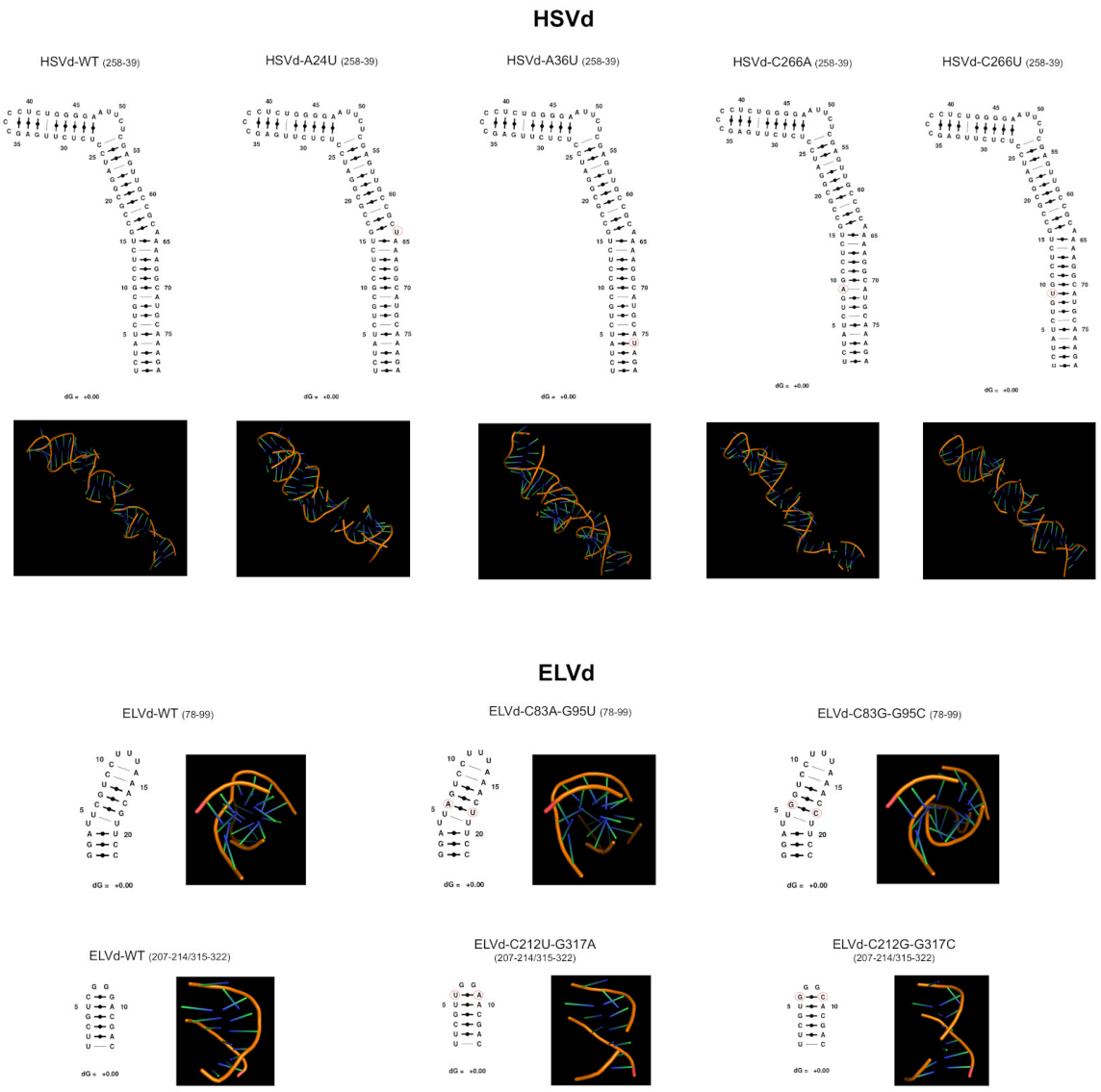
Supplementary Figure 2.4: E-ORF1 and E-ORF2 act as true ChTP. A) *In silico* prediction of chloroplast transport peptides in ELVd. Length: is the length of the submitted sequence. Score is the output score from the second step network. The prediction cTP/no cTP is based solely on this score. cTP tells whether or not this is predicted as a cTP-containing sequence; "Y" means that the sequence is predicted to contain a cTP; "-" means that is predicted not to contain a cTP. CS-score is the MEME scoring matrix score for the suggested cleavage site. **B**) The ORF-1 and E-ORF2 peptides are cleaved like canonical ChTP. Unprocessed E-ORF1 was not detected, indicating that this peptide is more efficiently cleaved. The detection of unprocessed E-ORF2 is consistent with the observation of chimeric E-ORF2/YFP in cell cytoplasm (Figure 2B).Total protein was extracted from infiltrated leaves and the E-ORFs/YFP peptides were analyzed by western blot assay with YFP-antisera. Not processed (np) and proccessed (p) peptides are marked with arrows. Protein extracts recovered from leaves infiltrated with unmodified YFP (UM/YFP) were used as a control. **C**) The predicted ChTP guided the specific compartmentalization of ELVd-derived peptides. Constructs carryng cTP-depleted E-ORF1 (E-ORF1/no-cTP) and E-ORF2 (E-ORF2/no-cTP) fussed to YFP were unable to be localized in chloroplasts and showed a localization pattern comparable to free-YFP. Scale bars are 10 µm.



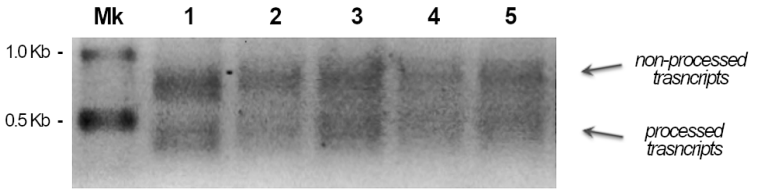
Supplementary Figure 2.5: Subcellular localization of the putative viroid-derived peptides in infected tissues. Confocal microscopy images of HSVd-infected *N. benthamiana* (**A**) and ELVd-infected eggplant (**B**) leaves agro-infiltrated with constructs expressing H-ORF3 fused to GFP and E-ORF1 in frame with YFP, respectively. The fusion fibrillarin-ChFP (central panels in A) was used as a nucleolus marker. Chloroplasts were determined by the chlorophyll auto-fluorescence recovered above 700 nm (central panels in B). Scale bars are 5 μm. The viroid accumulation (HSVd in A and ELVd in B) in infiltrated tissues was determined by northern blots assays (right panels). 1: Total RNA extracted from infiltrated leaves. 2: HSVd and ELVd monomeric transcripts of positive polarity used as control. *N. benthamiana* and eggplant leaves were inoculated by agro-infiltration with constructs expressing dimeric HSVd and ELVd transcripts respectively. The H-ORF3/GFP and E-ORF1/YFP constructs were agro-infiltrated in inoculated leaves at 21 days after inoculation. Viroid-inoculation, northern blot and transient expression assays were performed as is described in methods section.



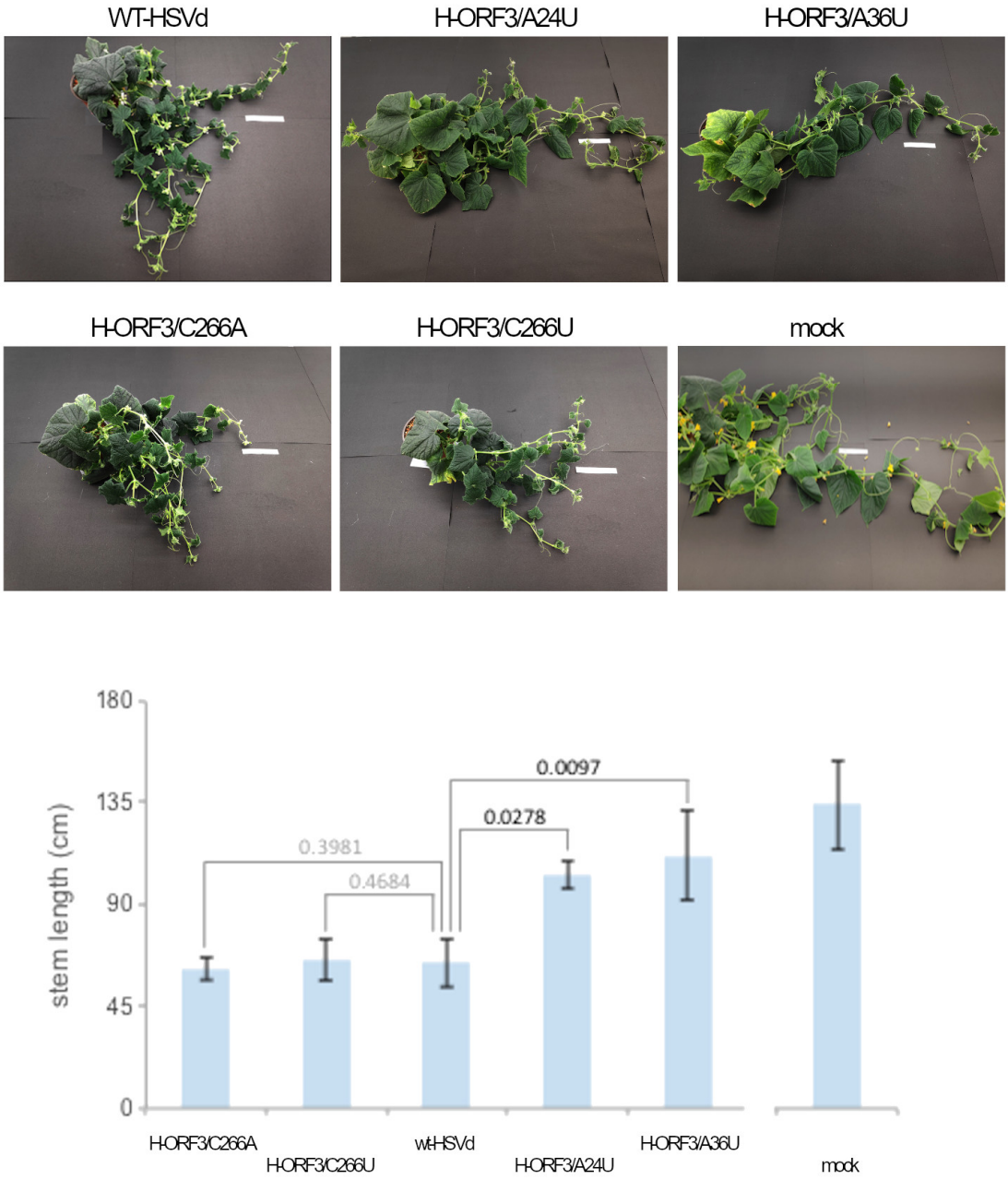
Supplementary Figure 2.6: Secondary structure of the HSVd and ELVd monomeric forms. General viewing of the secondary structures predicted for unmodified (wt) and mutant variants used in this work. The sequence modifications performed in each mutant are marked with arrows and painted in red. The procedure for prediction of the secondary structure is detailed in the method section. Minimal free energy (kcal/mol) of the predicted secondary structures was calculated with RNAeval web server at 25 ºC and 30 ºC for ELVd and HSVd sequences, respectively, and assuming RNA molecules to be circular.



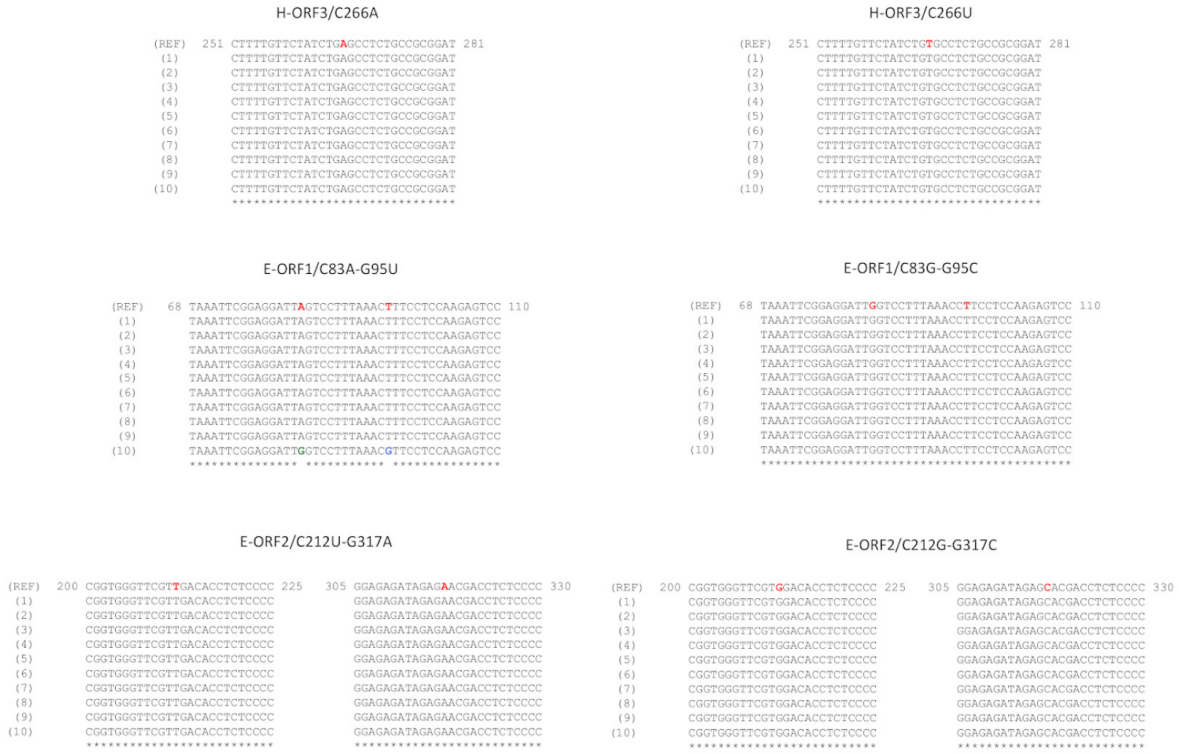
Supplementary Figure 2.7: Predicted RNA secondary and tertiary structures of the modified regions in viroid sequences. The 2D and 3D structures of HSVd are shown in the upper panel while those of ELVd in the lower panel. The RNA structures were generated with the MC-Fold and MC-Sym pipeline (Major et al., 2008. Nature 452: 51–55) using default parameters and considering H-type pseudoknots. The images of the RNA tertiary structure (right) were obtained using PyMol. The range of viroid sequence used for the predictions is shown in brackets. The modified nucleotides are marked with a red circle.



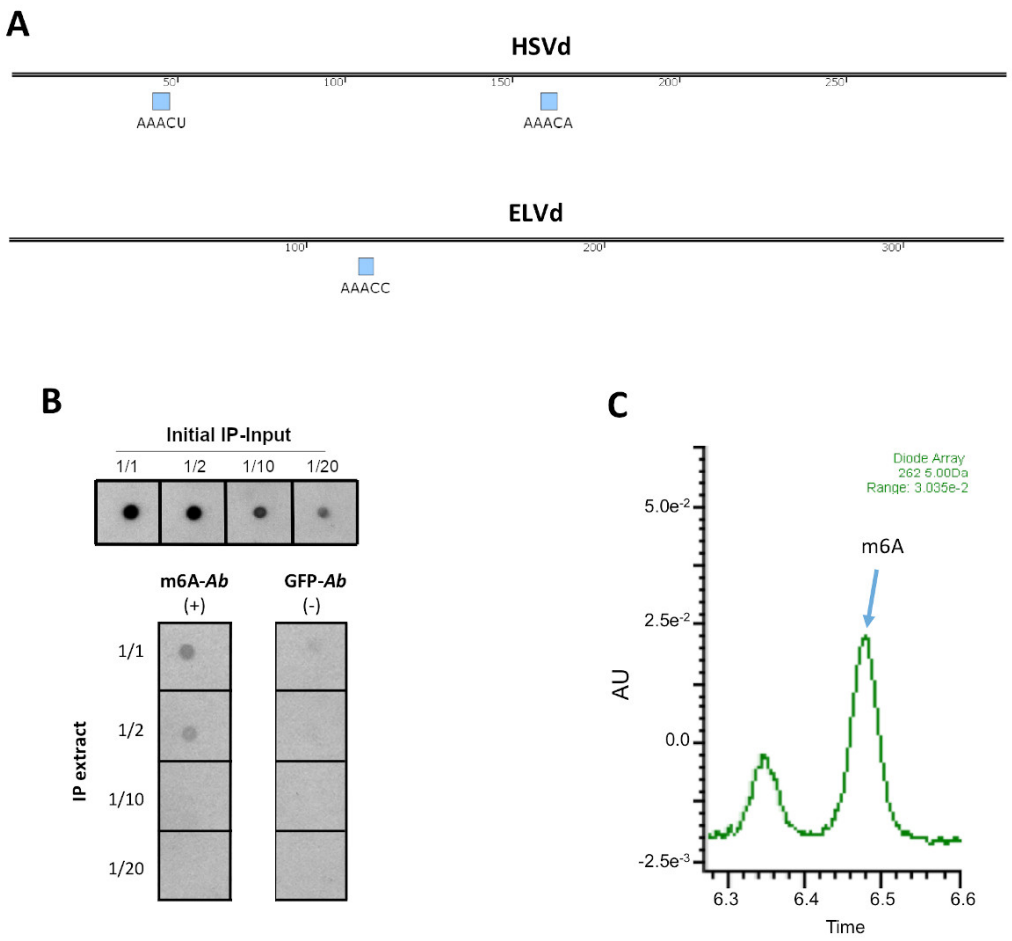
Supplementary Figure 2.8: Autocatalytic ribozyme activity is not altered in ELVd-mutants. Agarose gel showing an *in vitro* transcription of the dimeric cDNA of the unmodified ELVd (1) and E-ORF1/C83A-G95U (2), E-ORF1/C83G-G95C (3) and E-ORF2/C212U-G317A (4), E-ORF2/C212G-G317C (5) mutants. The transcription reaction was stopped after 5 minutes by adding 1V of inhibition solution (50 mM EDTA, 8M UREA and 50% (w/v) formamide) and then loaded into the agarose gel at 1%. Mk: ssRNA ladder.



Supplementary Figure 2.9: Symptom induction is impaired in plants infected with truncated-H-ORFs. A) Representative cucumber plants inoculated with unmodified (WT) and mutant HSVd variants or empty vector (mock). B) Graphic representation of the stem length in inoculated plants. The statistical significance was estimated by paired t-test and the obtained p-values are shown, not significant values are in grey. Error bars represent the standard error values.



Supplementary Figure 2.10: Nucleotide conservation of the mutant-viroid variants. Total RNA extracted from apical leaves of infected plants (at three weeks after inoculation) was amplified by RT-PCR with specific HSVd and ELVd primers and cloned into pJET1.2. Individual clones (10) were sequenced for each mutant variant, except for HSVd-A24U and HSVd-A36U which could not be amplified by RT-PCR. Alignments of the obtained sequences (1 to 10) with the sequence region flanking nucleotide substitutions in mutants variants (REF) are shown. Modified nucleotides in the mutant variants are in red. The unique nucleotide reversion to wild type sequence (G 95) and a substitution non-related to original sequence (G 83) in an individual clone of the E-ORF1/C83A-G95U variant, are marked in blue and green, respectively.



Supplementary Figure 2.11: N6-methyladenosine modifications in viroid RNAs. A) Canonical RRACH sites (being R [A or G] and H [A or C or U]) were computationally identified in HSVd (upper panel) and ELVd (lower panel). N6-methyladenosine (m6A) modifications predominantly occur in this RRACH context. **B**) Total RNA extracted from HSVd-infected cucumber plants was used for the RNA Immunoprecipitation (RIP) as is described in experimental procedures. The presence of HSVd in serial dilutions of initial IP-input extract was confirmed by dot blot hybridization (upper panel). Total RNA was inmunoprecipitated with with m6A antibody (m6A-Ab) and GFP-antibody (GFP-Ab) as negative control, respectively. RNAs immunoprecipitated with the m6A-Ab (+) or the negative GFP-Ab control (-), were diluted, blotted onto a nylon membrane and hybridized with the HSVd probe to detect the plus polarity of the viroid (lower panels). **C**) UPLC-PDA-Tof-MS analysis of purified circular forms of HSVd. For this analysis, circular forms of HSVd were isolated as previously described, digested into nucleosides and injected into a Micromass Q-TOF spectrometer coupled to an Acquity UPLC-PDA system (Waters) via an electrospray ionization (ESI) interface. The potential m6A peak (recovered at 6,48 seconds – blue arrow) was identified by comparison with the retention time of purified N6-methyladenosine (Sigma-Aldrich).

Supplementary Table 2.1: Detail of the primers used in this work

OLIGONUCLEOTIDES	SEQUENCE 5'-3'
Cloning to fluorescent proteins	
Fw fibrillarin Nb CATG	GGTCTCCCATGGTTGCACCAACTAGAGGTC
Rv fibrillarin Nb CTAG	GGTCTCGCTAGCGGCAGCAGCCTTCTGCTTC
Fw E-ORF1 CATG	GGTCTCCCATGGAGGGTGGGTGTACCTCT
Rv E-ORF1 CTAG	GGTCTCGCTAGGGGAACACATCCTTGAAAGGG
Fw E-ORF2 CATG	GGTCTCCCATGATTCGGAGGATTCGTCC
Rv E-ORF2 CTAG	GGTCTCGCTAGCTGGGGAGAGGTCGTCCTC
Fw E-ORF3 CATG	GGTCTCCCATGGTGTGGTTCGGCGAATG
Rv E-ORF3 CTAG	GGTCTCGCTAGCTCTCTCCTGGAAGGCCGG
Fw H-ORF1 CATG	GGTCTCCCATGAGGCTCTGCCTTCGAAACAC
Rv H-ORF1 CTAG	GGTCTCGCTAGCAGAGAGGATCCGCGGCAG
Fw H-ORF2 CATG	GGTCTCCCATGGCAGGGAGGTGCTTACC
Rv H-ORF2 CTAG	GGTCTCGCTAGCGTTTTTTTCTTTGCATGCC
Fw H-ORF3 CATG	GGTCTCCCATGGGGGAATTCTCGAGTTGC
Rv H-ORF3 CTAG	GGTCTCGCTAGCGAGAGGATCCGCGGCAGAG
Fw H-RK CATG	CATGCGCAAAAGGCATGCAAAGAAAAAAA
Rv H-RK CTAG	CTAGTTTTTTTCTTTGCATGCCTTTTGCG
Fw P-RK CATG	CATGAGAAAAGAAAAAAGAAGGCGGC
Rv P-RK CTAG	CTAGGCCGCCTTCTTTTTTCTTTTCT
Fw E-ORF1 No-cTP CATG/b	GGTCTCCCATGTGTGGTTCGGCGAATGTACCG
Rv E-ORF1 CTAG/b	GGTCTCGCTAGCGGGAACACATCCTTGAAAGGGG
Fw E-ORF2 No-cTP CATG/b	GGTCTCCCATGTGCTCCGGCCTTCCAGGAGA
Rv E-ORF2 CTAG/b	GGTCTCGCTAGCTGGGGAGAGGTCGTCCTC
Mutagenesis HSVd	
Fw H-ORF3/A24U	CTGGGGAATTCTCGAGTTGCCGCtAAAGGC
Fw H-ORF3/A36U	CTGGGGAATTCTCGAGTTGCCGCAAAAGGCATGCAtAGAAAAAAAC
Rv H-ORF3/C266A	AGGGGCTCAAGAGAGGATCCGCGGCAGAGGCtCAGATAGAAC
Rv H-ORF3/C266U	AGGGGCTCAAGAGAGGATCCGCGGCAGAGGCaCAGATAGAAC
Fw D1-HSVd	GGTCTCCCATGGCTGGGGAATTCTCGAGTTG
Rv D2-HSVd	GGTCTCAAGGGGCTCAAGAGAGGAT
Fw D3-HSVd	GGTCTCCCCCTGCTGGGGAATTCTCGAGTTG
Rv D4-HSVd	GGTCTCTCTAGAAGGGGCTCAAGAGAGGAT
Mutagenesis ELVd	
Fw E-ORF1/C83A-G95U	CTTTAAATTCGGAGGATTAGTCCTTTAAACTTTCCTCCAAGAGTCCC
Rv E-ORF1/C83A-G95U	GGGACTCTTGGAGGAAAGTTTAAAGGACTAATCCTCCGAATTTAAAG
Fw E-ORF1/C83G-G95C	CTTTAAATTCGGAGGATTGGTCCTTTAAACCTTCCTCCAAGAGTCCC
Rv E-ORF1/C83G-G95C	GGGACTCTTGGAGGAAGGTTTAAAGGACCAATCCTCCGAATTTAAAG
Fw E-ORF2/C212U	CGACGGTGGGTTCGTTGACACCTCTCCCCCTC
Rv E-ORF2/C212U	GAGGGGGAGAGGTGTCAACGAACCCACCGTCG
Fw E-ORF2/G317A	CAGGAGAGATAGAGAACGACCTCTCCCCA
Rv E-ORF2/G317A	TGGGGAGAGGTCGTTCTCTATCTCTCCTG
Fw E-ORF2/C212G	GACGGTGGGTTCGTGGACACCTCTCCCC
Rv E-ORF2/C212G	GGGGAGAGGTGTCCACGAACCCACCGTC
Fw E-ORF2/G317C	CAGGAGAGATAGAGCACGACCTCTCCCC
Rv E-ORF2/G317C	GGGGAGAGGTCGTGCTCTATCTCTCCTG
Colony PCR	
Fw 35S	AAGGGATGACGCACAATCCC
Rv Popit	TGGATGATCTCTTTCTCTTATTCAG
***In vitro* transcription**	
Fw T7 965	TAATACGACTCACTATAGGGGCTTCCAGGATCGGTACCATG
Rv 966	CCAGAAGATGGACAAGTCTAGGG
qPCR	
Fwq HSVd	CTCAGAATCCAGCGAGAGGC
Rvq HSVd	CTCAAGAGAGGATACGCGGC
FwHK ubiquitin cucumber	GATTTTCGCCGGAAAGCAGC
RvHK ubiquitin cucumber	TGTAGAACTGAAGGACGGCG
Fwq ELVd	GAGTCCCTTCCCCAAACCCT
Rvq ELVd	AGAGGTACACCCACCCTCC
FwHK Cyclophilin eggplant	GCGCCAAATTCAAGGACGAGAACT
RvHK Cyclophilin eggplant	ACAGCCTCGGCCTTCTTAATCACA
RT-PCR of polysome fractions	
Fw PSTVd	CGGAACTAAACTCGTGGTTCCTG
Rv PSTVd	AGGAACCAACTGCGGTTCC
Fw EF1α	ACTGTGCTGTCCTCATTATTG
Rv EF1α	AGGGTGAAAGCAAGAAGAGC
Fw Cyclophilin	GCGCCAAATTCAAGGACGAGAACT
Rv Cyclophilin	ACAGCCTCGGCCTTCTTAATCACA
Fw ELVd	GAGTCCCTTCCCCAAACCCT
Rv ELVd	GGGGAGAGGTGTCGACGAA
Fw HSVd	CTCAGAATCCAGCGAGAGGC
Rv HSVd	CTCAAGAGAGGATACGCGGC

CHAPTER 3

Integrative time-scale and multi-omic analysis of host-responses to *Hop stunt viroid* infection

Joan Marquez-Molins[1,2^], Pascual Villalba-Bermell[1^], Julia Corell-Sierra[1], Vicente Pallas[2] and Gustavo Gomez[1*].

[1] Institute for Integrative Systems Biology (I2SysBio), Consejo Superior de Investigaciones Científicas (CSIC) - Universitat de València (UV), Parc Científic, Cat. Agustín Escardino 9, 46980 Paterna, Spain.

[2] Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de Investigaciones Científicas (CSIC) - Universitat Politècnica de València, CPI 8E, Av. de los Naranjos s/n, 46022 Valencia, Spain.

[^] Equal contribution

[*] *Corresponding author*: E-mail: gustavo.gomez@csic.es

Preprint: https://www.biorxiv.org/content/10.1101/2022.01.06.475203v1

doi: 10.1101/2022.01.06.475203

PERSONAL CONTRIBUTION

I performed all the experiments, analyzed and discussed the results, and performed computational analysis.

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Economy and Competitiveness (co-supported by FEDER) Grants PID2019-104126RB-I00 (GG) and PID2020-115571RB-I00 (VP). I have been recipient of a predoctoral contract from the ACIF programme (ACIF-2017-114) of the Conselleria d'Educació, Investigació, Cultura i Esport Generalitat Valenciana.

Abstract and keywords

ABSTRACT

Constricted by an extreme biological simplicity, viroids are compelled to subvert host regulatory networks in order to accomplish their infectious process. Most of the studies focused on the response to viroid infection have only addressed a specific host regulatory level and considered a unique infection time. Thus, much remains to be done if we want to understand the temporal dynamics and complex nature of viroid-host interactions. Here we present an integrative analysis of the timing and intensity of the genome-wide alterations in cucumber plants infected with hop stunt viroid (HSVd). Differential host transcriptome, sRNAnome and methylome were integrated to determine the temporal response to viroid-infection. Our results support that HSVd promotes a dynamic redesign of the cucumber regulatory pathways predominantly affecting specific regulatory layers at different infection phases. Remarkably, the initial response was characterized by a reconfiguration of the host transcriptome by differential exon usage, followed by a predominant down-regulation of the transcriptional activity possibly modulated by the host epigenetic changes associated to infection and characterized by an increased hypermethylation. The alteration in host sRNA and miRNA metabolism was marginal. We expect that these data constituting the first comprehensive map of the cucumber-response to HSVd could contribute to elucidate the molecular basis of the host alterations triggered by viroid infection.

KEYWORDS

Viroid-host interactions, epigenetic epidemiology, host transcriptional regulation in response to pathogens, global response to biotic stress, systems biology and diseases.

Introduction

Viroids (the smallest infectious agents known to date) comprise an intriguing group of plant pathogens affecting a varied range of hosts worldwide. These sub-viral pathogens are classified into two families, *Pospiviroidae* whose replication takes place in the nucleus, and the *Avsunviroidae,* that replicates in the chloroplast (Di Serio et al., 2017). As naked circular viroids are compelled to closely interact with diverse (many yet unknown) host factors in order to fulfil their infectious cycle in the infected plants (Ding, 2009; Gómez and Pallás, 2013; Adkar-Purushothama and Perreault, 2020). However, how these pathogenic RNAs alter the host development and physiology inducing disease symptoms is still an unsolved question (Gago-Zachert, 2016; Adkar-Purushothama and Perreault, 2020; Navarro et al., 2021a).

In general, plants respond to stress conditions (including pathogen infection) through a complex reprogramming of their transcriptional activity. Gene regulation is a complex process modulated by a series of coordinated events involving multiple control layers such as: epigenetic modifications, modulation of the transcriptional activity and small RNA interference (Baulcombe and Dean, 2014; Yaish, 2017; Annacondia et al., 2018). These different regulatory levels have been categorized as stress-responsive structures termed as Environmental Gene Regulatory Influence Networks (EGRINs) (Wilkins et al., 2016).

Increasing evidence indicate that viroids can subvert different host-EGRINs and consequently promote the emergence of the phenotypic alterations recognized (in certain viroid-host interactions) as symptoms. In this sense, diverse studies have provided support to the notion that during the infectious process, viroids promote significant alterations in diverse plant regulatory pathways (Adkar-Purushothama and Perreault, 2020; Navarro et al., 2021a).

Although it has been proposed that viroids (presumably because of the highly compacted and structured genomic circular RNA), are resistant to degradation mediated by RNA-silencing (Gómez and Pallás, 2007a; Itaya et al., 2007; Elena et al., 2009), early reports evidenced that viroid-derived sRNAs (vd-sRNAs) (probably arising from the DCL-mediated processing of replication intermediates), are detected in plants infected by members of both nuclear and chloroplastic families (Itaya et al., 2001; Papaefthimiou et al., 2001; Martinez de Alba et al., 2002). The possibility that the vd-sRNAs could guide the silencing of plant-endogenous transcripts triggering the induction of disease symptoms, was then proposed as a plausible hypothesis to explain the basis of viroid-induced pathogenesis (Papaefthimiou et al., 2001; Wang et al., 2004; Gómez et al., 2009). Several research groups have demonstrated in the last years the involvement of vd-sRNAs in the

down-regulation of host transcripts (Wang et al., 2011; Navarro et al., 2012b; Adkar-Purushothama et al., 2015; Avina-Padilla et al., 2015; Adkar-Purushothama et al., 2017; Adkar-Purushothama et al., 2018; Adkar-Purushothama and Perreault, 2018; Bao et al., 2019; Navarro et al., 2021b), thus emphasizing the role of RNA silencing in viroid-induced symptoms. In addition, it has been recently proposed that viroid infection might trigger the production of secondary siRNAs able to recognize host transcripts that are not direct targets of the vd-sRNAs (Adkar-Purushothama et al., 2017). Also related to viroid-pathogenesis and RNA-silencing, alterations in the accumulation of certain endogenous miRNAs have been reported in diverse viroid-host interactions (Diermann et al., 2010; Owens et al., 2012; Tsushima et al., 2015; Sanz-Carbonell et al., 2019; Sanz-Carbonell et al., 2020; Apostolova et al., 2021; Dang et al., 2021; Navarro et al., 2021b).

Modulation (in both positive and/or negative sense) of the transcriptional activity constitutes the core of the plant response to abiotic and biotic stress conditions (Waters et al., 2017; Wang et al., 2020). Changes at the transcriptomic level have been described in several plants infected by viroids of both *Pospiviroidae* and *Avsunviroidae* families (Rizza et al., 2012; Herranz et al., 2013; Katsarou et al., 2016b; Kappagantu et al., 2017; Xia et al., 2017; Góra-Sochacka et al., 2019; Štajner et al., 2019). Altogether, these studies revealed a global alteration in the expression of host genes related to defense, response to stress, hormone homeostasis and signaling, biosynthesis of cell wall compounds, and RNA metabolism, among others (Adkar-Purushothama and Perreault, 2020).

A growing body of evidence points toward the existence of a close interplay between infection and epigenetic alterations in the host (plant or animal) genome (Annacondia et al., 2018; Wang et al., 2019a). Consequently, epigenetic epidemiology has emerged as a promising area for future research on infectious diseases (Gómez-Díaz et al., 2012). Recent studies have evidenced dynamic changes in host DNA methylation patterns (mainly rDNA and transposable elements) upon viroid infection in both somatic (Martinez et al., 2014; Castellano et al., 2016b) and reproductive (Castellano et al., 2016a) tissues. The changes of the host-genome epigenetic regulation have been connected (in infected cucumber plants) with viroid recruiting and functionally subverting the host HISTONE DEACETYLASE 6 (HDA6) (Castellano et al., 2016b). Alteration of the host plant methylome has been also described to occur in response to infection by viruses (Corrêa et al., 2020) and bacteria (Deleris et al., 2016), supporting that regulation of host DNA methylation may be part of an evolutionary conserved immune response in plants (Zhu et al., 2016).

The exposed above evidenced that the advent of the high throughput technologies has favored the study of the viroid-host interactions at a molecular level. However, such experimental approaches have been generally centered on the analysis of the viroid-induced changes at a single host-regulatory level and/or considering a particular infection time (mainly correlated to advanced

disease stages). Therefore, our knowledge about the functional diversity and temporal dynamics of the global plant response to viroid infection is scarce (Adkar-Purushothama and Perreault, 2020).

Here we have performed an integrative analysis of the timing and intensity of the genome-wide alterations in cucumber (*Cucumis sativus*) plants in response to hop stunt viroid (HSVd) infection (Marquez-Molins et al., 2021). Our study has been focused on the temporal dynamics (at three infection time-points) of the plant response to viroid infection at three different regulatory levels of the gene expression network: *i)* small RNA interference, *ii)* modulation of the transcriptional activity and *iii)* epigenetic modifications. The obtained results evidenced that HSVd infection promotes a dynamic redesign of the cucumber regulatory pathways that affects these regulatory layers distinctively depending on the infection phase.

Results

Viroid infection

To uncover the dynamic of the host regulatory response to HSVd in cucumber we analyzed apical leaves of infected plants at 10-, 17- and 24-days post inoculation (dpi). Mock-inoculated plants were used as control. Time points were arbitrarily selected as representative of the early, intermediate and late HSVd infection phases, respectively. As expected, viroid induced symptoms were evident in HSVd-infected plants at 24 dpi (Figure 3.1A). A total of 54 libraries (18 for each sRNAs, long transcripts and methyl-C analysis) consisting of three bio replicates per time point of infected and control plants were obtained (Supplementary Figure 3.1). The information about the number and quality of the obtained reads is detailed in the Supplementary Table 3.2 for each omic approach respectively. Sequences used in this study are available at NCBI repository (BioProject ID: PRJNA795883).

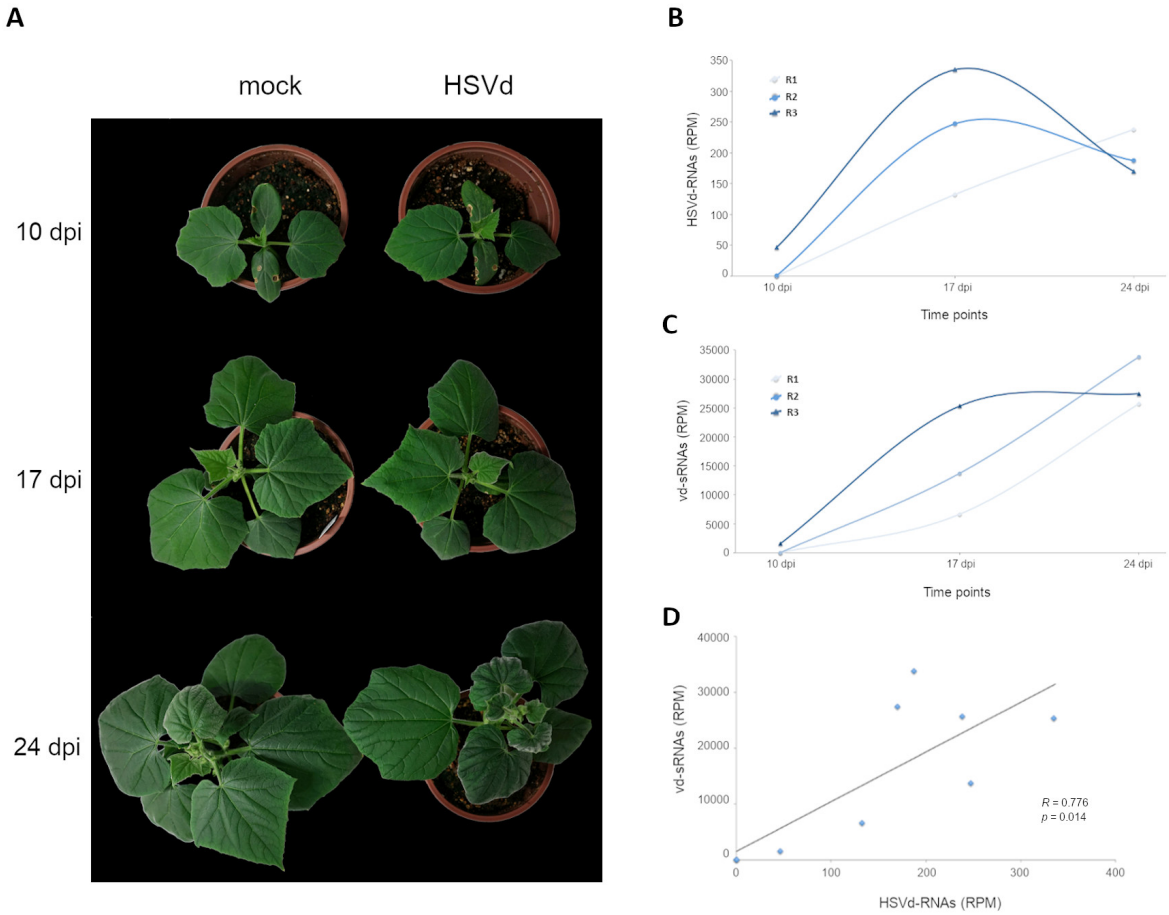


Figure 3.1. HSVd RNA is detected at 10 dpi in infected plants: A) Representative infected and mock inoculated cucumber plants at the three analyzed time points. Typical plant symptoms characterized by reduction in leaf size and incipient stunting are evident at 24 dpi. Graphic representation of the total transcripts (**B**) and sRNAs (**C**) derived from HSVd genome detected in apical leaves at 10, 17 and 24 dpi. **D**) Scatter plot showing the significant positive correlation (estimated by Pearson correlation coefficient) between the temporal accumulation of vd-transcript and viroid sRNAs (vd-sRNAs) in infected cucumber plants.

The infection rate of inoculated plants was estimated by considering the accumulation of viroid transcripts and viroid-derived sRNAs (vd-sRNAs) (Figure 3.1B and C). Viroid transcripts were detected in two (HSVd-R2 – 0.53 RPM and HSVd-R3 – 46.4 RPM) of the three analyzed samples at 10 dpi, while at 17 dpi an increased number of HSVd-related transcripts were recovered from all the analyzed samples (132.48, 247.13 and 335.04 RPM for replicates HSVd-R1, HSVd-R2 and HSVd-R3 respectively) (Supplementary Table 3.3). A lower accumulation (in comparison to 17 dpi) of viroid-transcripts was observed in HSVd-R2 (187.3 RPM) and HSVd-R3 (169.4 RPM) samples at 24 dpi. In contrast, the biological replicate HSVd-R1 (with no-detected viroid transcripts at 10 dpi), showed an increased accumulation (238.0 RPM) at 24 dpi (Supplementary Table 3.3). On the other hand, vd-sRNAs were only detected (1421.0 RPMs) in systemic tissues of the replicate HSVd-R3 at 10 dpi. The accumulation of sRNAs homologous to HSVd in infected tissues at intermediate and late HSVd infection phases was more homogenous and vd-sRNAs were recovered from the three analyzed replicates at 17 dpi (6456.6 RPM, 13504.1 RPM and 25128.6 RPM) and 24 dpi (25539.9 RPM, 33627.7 RPM and 27284.7 RPM), revealing that vd-sRNAs increased through the analyzed period (Figure 3.1C). In general, vd-sRNAs accumulation correlated ($R = 0.776$, $p = 0.014$) with the recovered HSVd transcripts (Figure 3.1D). HSVd transcripts and vd-sRNAs represented (in average) respectively, the 0.002% and 0.047% of the recovered reads at 10 dpi, the 0.024% and 1.5% at 17 dpi and the 0.02% and 2.9% at 24 dpi (Supplementary Table 3.3 and 3.4). Our results support that both, viroid transcripts and vd-sRNAs may be detected in systemic tissues in an irregular manner in early infection phases (10 dpi) and more consistently since the 17 dpi.

Modifications in the host siRNA/miRNA population

Associations between sRNA expression profiles (considering control/infected plants and their biological replicates in the three time points) were evaluated using Principal Components Analysis (PCA). The percentages of variance explained by the first three principal components (PC) were 16.43%, 15.01% and 14.47%, respectively (45,91% of the total variance). The biological replicates clustered together (attesting for the reproducibility of our assays) and treatments/time-points groups were significantly ($p = 3.05 \times 10^{-12}$) separated in the PC space (Figure 3.2A).

The endogenous cucumber sRNAs exhibited a distribution of read lengths enriched for 24 nt long (34.0%), followed by 23 nt (23.2%) and similar accumulations of 25 (14.3%), 21 (12.0%) and 22 (10.5%) nt long molecules. Reads of 20 nt represented the least abundant category (6.0%) (Figure 3.2B). This difference in accumulation of different sRNA lengths was statistically significant (2-ways non-parametric ANOVA, $p < 2.12 \times 10^{-15}$, 2.14×10^{-13} and 1.30×10^{-15} for 10, 17 and 24 dpi (respectively). In coincidence with diverse cucumber cultivars (Ling et al., 2017; Xu et al., 2019),

this effect was predominantly due to the large enrichment in 24 nt long sRNAs in all pairwise comparisons. No significant differences were found between HSVd-infected and control conditions regarding the observed global distribution of sRNAs sizes (2-ways non-parametric ANOVA, p = 0.93; 0.85 and 0.55 for 10, 17 and 24 dpi, respectively).

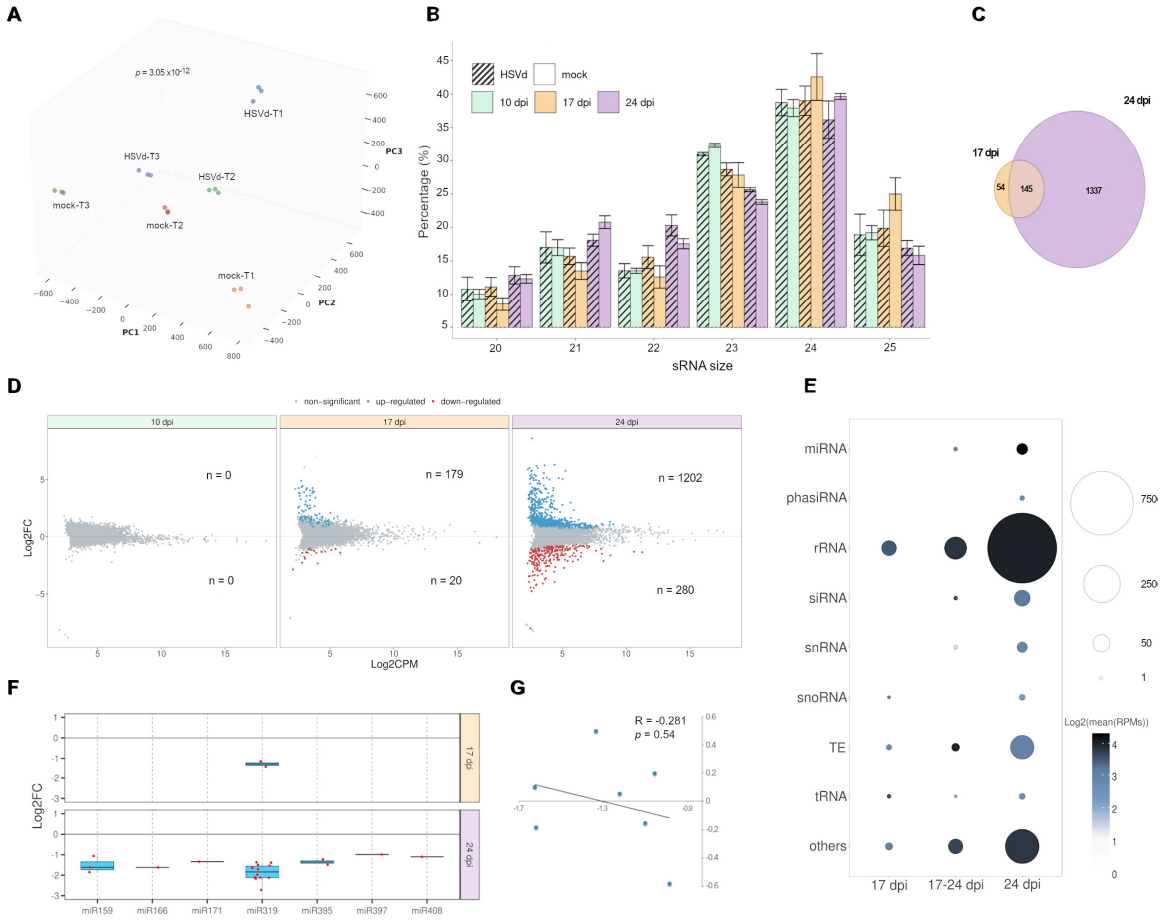


Figure 3.2. Analysis of the sRNA population recovered from the analyzed cucumber libraries. A) Principal component analysis based on sRNAs accumulation in biological replicates of HSVd-infected and mock-inoculated plants at the three analyzed time-points. The statistical significance was estimated by Mann-Whitney-Wilcoxon test, considering the inter- and intra-group Euclidean distances. **B**) Diagram showing the means of the temporal relative accumulation and distribution of the total sRNAs reads ranging between 20 and 25 nt (Error bars indicate the SE between biological replicates). The three analyzed times are represented by colors (green: 10 dpi, orange: 17 dpi and magenta: 24 dpi). Smooth and striped bars represent data arising from mock-inoculated and HSVd-infected plants. **C**) Venn diagram representing the number of sRNAs responsive to HSVd infection at 17 and 24 dpi. **D**) Volcano plot representing the endogenous cucumber sRNAs with significant differential expression in infected plants at the three analyzed time points. Blue and red dots indicate up- and down- regulated sRNAs, respectively. The number of differential sRNAs in each time point is detailed. **E**) Categorization of HSVd-responsive sRNAs. The ball size represents the number of unique reads recovered. The color intensity indicates the sRNA accumulation estimated by the Log2 of the mean of the normalized reads. **F**) Box-plot analysis showing the general expression value observed for each miRNA-family member (dots) at 17 and 24 dpi. The differential expression of each miRNA family is represented by the median (internal box line) of the LFC values. **G**) Scatter plot showing the non-significant negative correlation (estimated by Pearson correlation coefficient) between the expression levels of the seven miRNAs responsive to HSVd infection and the accumulation of their targets in infected plants.

Only a reduced number (1536 unique reads) of endogenous sRNAs showed significant altered accumulation in HSVd-infected plants at 17 and 24 dpi (no differential sRNAs were detected at 10 dpi) (Figure 3.2C). 145 sRNAs were differentially expressed at both 17 and 24 dpi, while 54 and 1337 endogenous sRNAs, showed specific differential accumulation at both 17 and 24 dpi, respectively. A total of 1246 (81,12%) and 290 (18.88%) sRNAs increased or decreased, respectively, their accumulation in response to infection. Under a temporal viewpoint, although no differentially expressed sRNAs were detected at 10 dpi, 199 reactive sRNAs (179 over-expressed and 20 down-regulated) were identified at 17 dpi (Figure 3.2D). The most evident alteration in the accumulation of endogenous sRNAs was observed at 24 dpi in which 1482 reads (1202 up-regulated and 280 down-regulated) were identified as differentially expressed in HSVd-infected plants. The 68.42% (1051 reads) of the differential sRNAs were categorized as derived from ribosomal RNA (r-sRNAs), while the 7.55% (116 reads) from Transposable Elements (TE-sRNAs), and the 3.06% corresponded to 47 reads identified as small interfering RNAs (siRNAs) previously described in cucumber (Figure 3.2E). A minor representation (1.5%) was observed for known microRNAs (miRNAs), sRNAs derived from snRNAs (1.3%), tRNAs (0.52%), snoRNAs (0.39%) and phased siRNAs (0.2%). The 17.06% of the sRNAs with differential expression were derived from unidentified or not representative functional categories.

We identified seven miRNA families (miR158, miR166, miR171, miR319, miR395, miR397 and miR408) with differential expression associated to HSVd-infection (Figure 3.2F). Except for miR319 (with 14 different sequences), reactive families were represented by one (miR166, miR171, miR397 and miR408), two (miR395) or three (miR159) unique sequences. All members in each miRNA family decreased in a coordinated manner in response to HSVd-infection (Supplementary Table 3.5). miR319 was the unique family that exhibit differentially expressed members at 17 and 24 dpi, the rest of the reactive miRNAs were only differential at 24 dpi. The differential expression values of the reactive miRNAs were low (in any case LFC > -1,6), suggesting that the regulatory pathways mediated by miRNAs are not particularly affected by viroid infection. The observation that such miRNAs negatively correlated in a non-significant manner ($R = -0.281$, $p = 0.541$) with their target transcripts reinforced this notion (Figure 3.2G).

Biogenesis and functional role of HSVd-derived sRNAs

A total of 1,523,822 reads (representing 7,428 unique sequences) fulfilling the conditions (detailed in material and methods) required to be unequivocally considered as HSVd-derived sRNAs (vd-sRNAs) were recovered from the three analyzed time points (Supplementary Table 3.3). Analysis of polarity distribution revealed a slight but significant ($p = 2.2x10^{-16}$) difference according to the strand of origin of the vd-sRNAs (58.1% sense and 41.9% antisense) (Figure 3.3A). This polarity proportion was maintained constant in the three analyzed time points (Supplementary Figure 3.2). Considering each sRNA size-class individually, the sense-strand derived sRNAs were the predominant, except in 20 nt and 23 nt in length vd-sRNAs (Figure 3.3B). In contrast, a comparable accumulation rate for both plus and minus reads was observed when unique vd-sRNAs were considered (Figure 3.3C). Vd-sRNAs were mainly of 24 nt (68.9%, 57.1% and 49.0%,) and 21 nt (15.5%, 24.1% and 30.0%), for 10, 17 and 24 dpi, respectively. Except for 22 nt (10.3%, 11.8% and 13.2%), vd-sRNAs of 20, 23 and 25 nt accumulated under 5% of the total sRNAs homologous to HSVd in the three analyzed time points (Figure 3.3D). Regarding the predominant vd-sRNAs (24 nt and 21 nt) it was evident that while the proportion of 21 nt vd-sRNAs increases during viroid-infection, the accumulation of 24 nt reads recovered from infected plants decreases during the analyzed period. The observation that a comparable size-evolution pattern was not detected for endogenous sRNAs (Supplementary Figure 3.3) supports that this significant (2-ways non-parametric ANCOVA, $p < 0.031$ and 0.039, for 21 and 24 nt respectively) temporal bias in the accumulation of this type of vd-sRNAs is specific for the HSVd-RNA and not related to the cucumber sRNAs biogenesis pathway. Categorized according to their 5' terminal nucleotide, vd-sRNAs with a C-end were the most abundant (43,8%), followed by reads with T- (24,6), G- (17,7) and A- (13,9) 5′ ends (Figure 3.3E). Similar distribution of 5-ends was observed when total vd-sRNAs were analyzed considering size and temporal dynamics (Figure 3.3F).

When total vd-sRNAs recovered at the three time points were plotted onto HSVd-genome we observed that in general, sense and antisense vd-sRNAs spreading along the entire HSVd genome showed a heterogeneous distribution pattern with four evident hyper accumulation peaks in the position 130 and 254 and -78 and -184, in genomic and anti-genomic HSVd-RNA, respectively (Figure 3.3G). The hyper accumulation peaks predominantly corresponded to 24 nt sRNAs (Figure 3H). This global profile was maintained during the analyzed infection period (Supplementary Figure 3.4). In contrast to hyper accumulation peaks, we observed that a well-defined region of the viroid-genome comprised between the positions 11 to 38 (sense reads) and 31 to 58 (antisense reads) showed a null or much lower accumulation of vd-sRNAs in the three analyzed time points (Supplementary Figure 3.5).

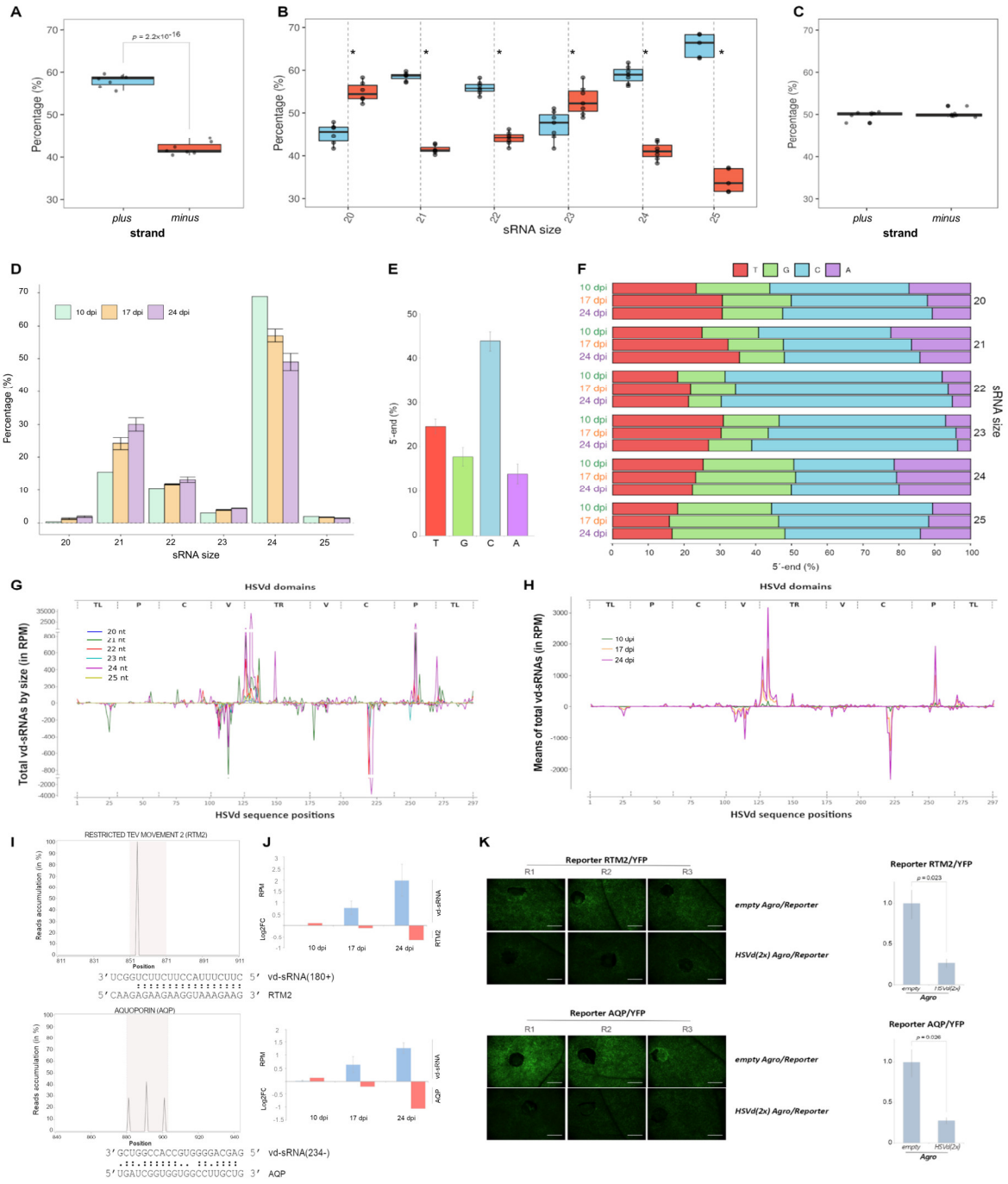


Figure 3.3. Characterization of vd-sRNAs. **A**) Box plot analysis of the relative accumulation of the total vd-sRNAs (20 to 25 nt) derived from genomic (plus) and antigenomic (minus) HSVd-RNA strand at the three analyzed time points. **B**) Distribution of polarity of the total vd-sRNAs discriminated by size, an asterisk indicates a p-value under 0.05. **C**) Comparative analysis of unique reads of vd-sRNAs. **D**) Diagram showing the means of the temporal relative accumulation and distribution of the vd-sRNAs reads ranging between 20 and 25 nts (the error bars indicate the SE). The three analyzed times are represented by colors (green: 10 dpi, orange: 17 dpi and magenta: 24 dpi). **E**) Graphic representation of the proportion of the 5'- ends in the total recovered vd-sRNAs. The means of the total reads are showed. **F**) Distribution of the 5'- ends discriminated by size and infection time. **G**) Genome view of the vd-sRNAs recovered from infected cucumber plants. The vd-sRNAs were plotted according to the position of their 5´-end onto the HSVd- RNA sequence in either sense (above the x-axis) or antisense (below the x-axis) configuration. The values on the y-axis represent the abundance of total vd-sRNAs (sizes indicated by colors) in the three analyzed time point. **H**) Representation of the temporal accumulation of total vd-sRNAs (means of total vd-sRNAs are represented). Vd-sRNAs recovered in each of the three analyzed times are represented by colors (green: 10 dpi, orange: 17 dpi and magenta: 24 dpi). **I**) Representation of the vd-sRNA-cleaved transcripts (RTM2 and AQP, upper and lower panel

respectively) differentially detected by degradome assay in infected plants. Degradome sequences were plotted onto the predicted targets transcripts. The grey zone indicates the predicted vd-sRNA recognition site (detailed below). The values on the Y-axis represent the relative accumulation of the reads matching onto the 100 nt in length analyzed region. **J**) Dynamics of the accumulation of each pair of vd-sRNA (blue) and its target (red) over the infection. Vd-sRNA accumulation is expressed as the mean of the reads per million (RPM) and the error bars indicate the SE between biological replicates. The accumulation of the predicted target transcript is represented as the Log2FC value. **K**) Down regulation of the predicted targets mediated by HSVd was validated by transient expression of specific reporters carrying YFP transcriptionally fussed to the RTM2 (upper panel) and AQP (lower panel) target sequences. Reporter expression was determined by fluorescence emission (left) and transcript quantification by RT-qPCR (right). Error bars indicate the SE between biological replicates.

It has been previously demonstrated that certain vd-sRNAs can trigger the silencing of endogenous transcripts in infected plants (Navarro et al., 2012b; Adkar-Purushothama et al., 2015; Adkar-Purushothama et al., 2017). Although some mRNAs have been proposed to be targeted by HSVd-derived sRNAs, the validation of this activity is still lacking for this viroid (Zhang et al., 2020). Therefore, we performed a degradome analysis starting from the RNA samples obtained (by triplicated) from the mock- and HSVd-inoculated plants at 24 dpi. According to our established parameters (Supplementary Figure 3.6) we found three vd-sRNAs predicted to target three cucumber transcripts with potential cleavage-site identified by degradome analysis and significant decreased expression in infected plants at 24 dpi suitable to be regulated by vd-sRNAs via RNA-silencing mechanisms (Figure 3.3I and Supplementary Figure 3.6).

The vd-sRNAs identified as vd-sRNA (180+), vd-sRNA (185-) and vd-sRNA (234-) were 21 nt in length and derived from *plus* (180+) and *minus* (185- and 234-) HSVd strands. The potential targets for vd-sRNAs were predicted to encode cucumber homologous to RESTRICTED TEV MOVEMENT 2 (RTM2), CELLULOSE SYNTHASE (*Ces*A) and AQUAPORIN (AQP). The relative accumulation of these predicted target-transcripts was consistently decreased through the infection period in contrast to the three vd-sRNA that increase their accumulation over time (Figure 3I and Supplementary Figure S7), reinforcing the data obtained by degradome analysis. An additional assay performed in *N. benthamiana* plants accumulating HSVd was used to validate the viroid-mediated down regulation of RTM2 and AQP by analysing the fluoresce emission and transcript accumulation of chimeric reporters containing the Yellow Fluorescent Protein (YFP) transcriptionally fussed to the transcript sequences predicted to be targeted by vd-sRNAs (180+) and (234-), respectively (Figure 3.3H and I).

HSVd-infection modulates the expression and alternative splicing of host-transcripts

It is expected that the host transcriptome may change in response to the viroid infection. To test that assumption, RNA extracted from control and infected plants was subjected to RNA-seq. A high proportion (98.33%) of the recovered transcripts in both infected and control plants mapped to the cucumber genome (Supplementary Table 3.3), allowing the detection of differentially expressed transcripts (DETs) in all time points. The residual reads corresponded to unassigned transcripts (1.66%) and HSVd-derived sequences (0.01%). Considering only reads fully homologous to previously annotated cucumber RNAs, we identified 1125, 515 and 1390 transcripts with significant (FDR ≤ 0.05) differential accumulation in HSVd-infected plants, at 10, 17 and 24 dpi, respectively (Figure 3.4A and Supplementary Table 3.6). A total of 925 (at 10 dpi), 288 (at 17 dpi) and 1052 (at 24 dpi) host-transcripts presented time-specific differential accumulation (Figure 3.4B). Considering Log2FC ≥ 1 or ≤ -1 as cut-off value, the list of differential transcripts was reduced to 10 up-regulated at 10 dpi, 55 (20 up-regulated and 35 down-regulated) at 17 dpi and 271 (46 up-regulated and 225 down-regulated) at 24 dpi (size-increased dots in Figure 3.4A). The 10 up-regulated genes recovered at 10 dpi were time-specific, whereas 27 and 243 transcripts showed specific differential expression at 17 and 24 dpi, respectively (Figure 3.4C). Under a temporal viewing, it was obvious that the host transcripts with increased accumulation were predominantly (695 transcripts, 61.7%) recovered at 10 dpi, while down-regulation was predominant at 17 dpi (302 transcripts, 58.64%) and 24 dpi (975 transcripts, 70.14%). A similar trend was observed when expression values with a Log2FC ≥ 1 or ≤ -1 were considered. The amplitude (estimated by the variance of the differential expression values) and diversity (estimated by the number of DETs) of the transcriptional response to HSVd was clearly increased during the analyzed period of infection (Figure 3.4D).

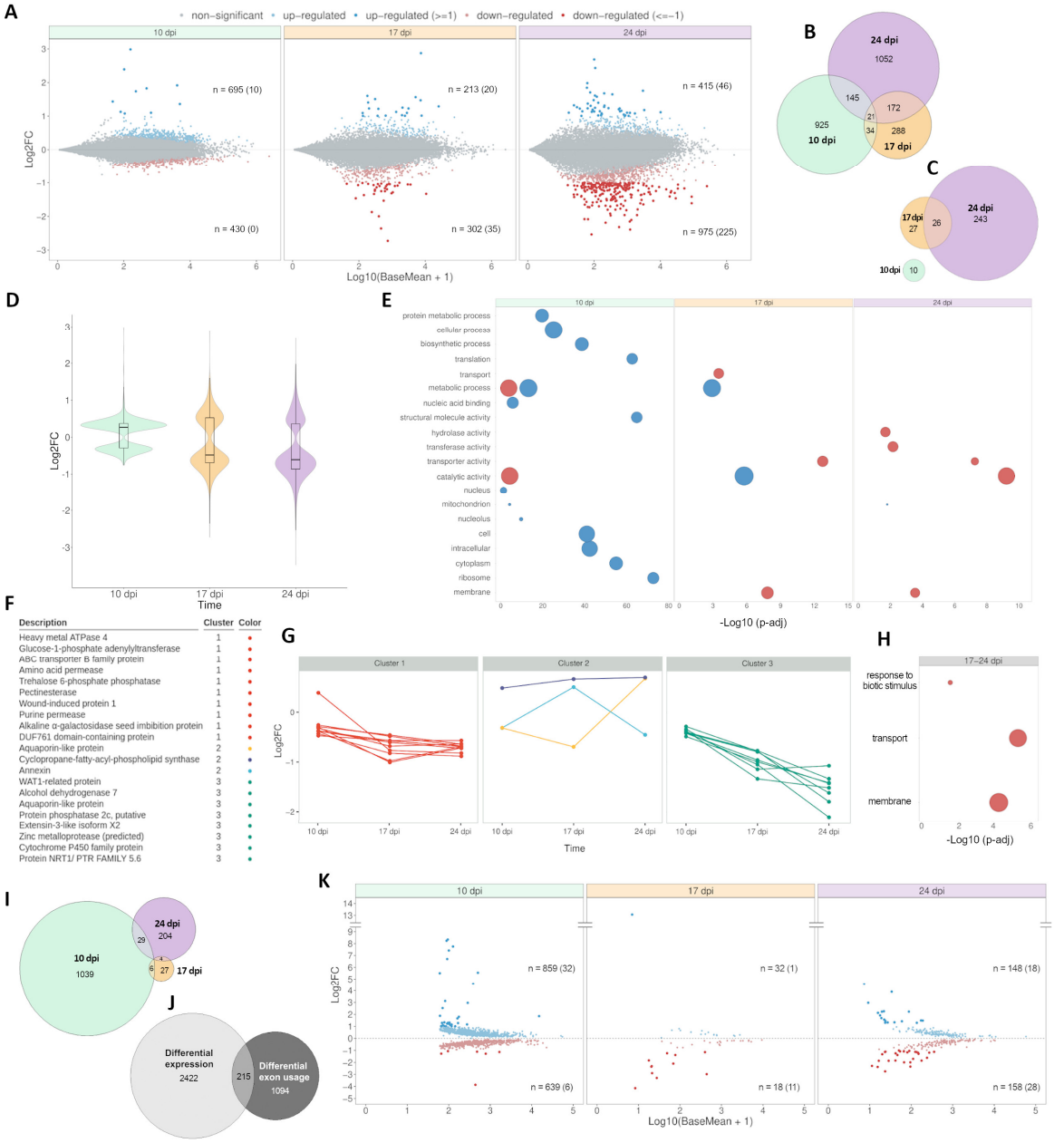


Figure 3.4. Host-transcriptional alterations associated to HSVd-infection. A) Volcano plot representing the differentially expressed transcripts (DET) in infected cucumber plants at 10, 17 and 24 dpi. Blue and red dots indicate up- and down- regulated transcripts, respectively. DET with LFC ≥ 1 or ≤ -1 are represented by bold dots and the number is indicated in brackets. Venn diagram showing all the DET (**B**) and those with LFC ≥ 1 or ≤ -1 (**C**) identified at 10, 17 and 24 dpi **D**) Violin plot representing the temporal profile of the transcriptional alteration associated to HSVd-infection. Internal box line indicates the median of the LFC values. **E**) Gene ontology analysis (plant GOSlim) for DETs identified in HSVd-infected plants at 10, 17 and 24 dpi. Circle size is proportional to the fraction of genes relative to the total number of genes with the GO term. Color indicates whether the genes are upregulated (blue) or downregulated (red). The –Log(10) of the adjusted P values is represented in the X-axis. **F**) Detail of the 21 transcripts differentially expressed at the three analyzed time-points. **G**) Clustering analysis of time-course expression profiling of transcripts with differential expression at 10, 17 and 24 dpi. **H**) Gene ontology analysis (plant GOSlim) for transcripts differentially expressed at both 17 and 24 dpi (LFC ≥ 1 or ≤ -1). **I**) Venn diagram showing the number of DEU events identified at 10, 17 and 24 dpi and (**J**) transcripts showing common and specific regulation by differential expression or exon usage. **K**) Volcano plot representing the events of differential exon usage (DEU) in infected plants at 10, 17 and 24 dpi. Blue and red dots indicate up- and down- regulated exons, respectively. Events of DEU with LFC ≥ 1 or ≤ -1 are represented by bold dots and the number is indicated in brackets.

According to the analysis of GO-terms, transcriptional response to HSVd-infection was characterized by a significant enrichment in genes associated to diverse biological categories mainly related to metabolic and cellular process, catalytic activity, ribosome metabolism, cellular/intracellular components, transport and membrane (Figure 4E and Supplementary Table 3.6). Under a global viewpoint, it was evident that the wide-range of enriched functional categories observed at 10 dpi (15 categories), was decreased during the infection (6 enriched-categories at 24 dpi). Searching for host genes potentially involved in a more persistent response to HSVd-infection we identified cucumber transcripts differentially expressed during the three infection phases. A total of 21 transcripts were identified as significantly responsive to infection at 10, 17 and 24 dpi (Figure 3.4F and Supplementary Table 3.7). Except for three specific transcripts up-regulated in at least one of the analyzed time points (Figure 3.4G, cluster 2), the common response of the other 18 differential transcripts to HSVd-infection was a stable (cluster 1) or temporally increased (cluster 3) down regulation. Genes included in this common response were functionally related to transport, membrane activity and response to biotic stimulus. This functional specialization was also supported by GO-terms enrichment when were considered only transcripts with Log2FC ≥ 1 or ≤ -1 with common differential expression at late (17 and 24 dpi) infection phases (Figure 4H). Differential transcripts without previous annotation (and consequently not considered in our study) were analyzed as is described in Methods section (Supplementary Figure 3.8). The obtained results revealed that such transcripts (mainly identified as non-coding RNAs) exhibited minimal alterations in response to HSVd infection (Supplementary Table 3.8).

Alternative splicing (AS) has been proposed as a regulatory mechanism crucial for the modulation of the plant development and response to virus infection (Mandadi and Scholthof, 2015). Sequenced transcripts were analyzed with DEXSeq tool to infer the AS landscapes of cucumber plants and determine their differential patterns during HSVd infection. We identified 1589 differential exons derived from 1309 intron-containing multiexonic genes that were alternatively spliced in infected plants (Figure 3.4I). Among such transcripts, only 215 (5.76 %) also showed differential expression (Figure 3.4J). Under a global viewpoint, significant differential exon usage (DEU) was predominantly observed in infected plants at 10 dpi (1498 exons with DEU), less frequent at 24 dpi (306) and a residual phenomenon at 17 dpi (with 50 alternative splicing events identified) (Figure 3.4K and Supplementary Table 3.9). Differential exon usage in response to HSVd-infection was comparable at early (38 exons) and late infection phases (46 exons) when we considered only Log2FC values ≥ 1 or ≤ -1 (size increased dots in Figure 3.4K). Regarding the temporal trend in the differential exons usage, our results support that HSVd infection is associated to a predominant retention of exons at 10 dpi showing 32 over-represented exons with Log2FC values ≥ 1 (Figure 4K). In contrast, the underrepresented exons (with Log2FC ≥ -1) were prevalent

at 17 (11 exons) and 24 dpi (28 exons). Transcripts exhibiting frequent DEU, at both early (10 dpi) and late (24 dpi) infection phases were predominantly involved in primary metabolism-related process (Supplementary Figure 3.9). Additionally, transcripts associated to membrane and transport, were enriched at 10 dpi.

Epigenetic alterations associated to HSVd-infection

To analyze if HSVd-infection modifies the epigenetic landscape of cucumber plants, whole-genome bisulfite sequencing (WGBS) libraries were constructed. Global analysis (considering the total percentage of cytosine methylation in CG, CHG and CHH sequence context) revealed a non-significant hypomethylation in infected plants at 10 dpi (70.39% and 67.21%, for control and infected respectively), followed by a slight hypermethylation at 17 dpi (75.04% and 76.93%) that was significant (70.6% and 73.21% for control and infected plants, respectively) at 24 dpi (Figure 3.5A, upper panel). These differences were maintained when we considered total methylation level for each sequence context specifically (Figure 3.5A, lower panel). A similar scenario was observed when the total methylation levels were analyzed, considering each cucumber chromosome individually (Figure 3.5B). To obtain a more precise picture of the specific methylation changes, we used DMRcaller to identify differentially methylated regions (DMRs) for the three CG, CHG, and CHH sequence contexts. Only significant DMRs with differences in methylation ≥15% were considered. The hypomethylated CG DMRs were the most abundant during the analyzed infection period. In contrast hypermethylated DMRs were most abundant considering CHG and CHH sequence context (Figure 3.5C).

This global trend was maintained constant during the three analyzed time points. Considering the intensity (assumed as the median of the DMRs values) of the methylation alterations in each time and context, it was evident that the higher differential values for both hypo (18.4%, 18.0% and 18.8% for 10, 17 and 24 dpi, respectively) and hypermethylation (19.8%, 17.5% and 18.3%10, 17 and 24 dpi, respectively) were observed for the GHG context, while a lower number of alterations was observed for the CHH DMRs (11.3%, 9.0% and 10.8%) for loss of methylation at 10, 17 and 24 dpi, respectively and 12.5%, 9.9% and 11.6% for gaining in identical time points (Figure 3.5D). For the DMRs of the CG context, the medians of the differential alterations were near to the 15% during the entire analyzed period. Considering the functional categories of the genomic position of the DMRs, approximately 60% of the observed CG DMRs in the three analyzed time points were mapped in regions predicted as coding sequences (CDS) and introns, while intronic regions constituted the predominant DMRs at the CHG context (Figure 3.5E). As expected, during the analyzed period, CHH DMRs were enriched in TEs-associated regions and strongly diminished in CDS.

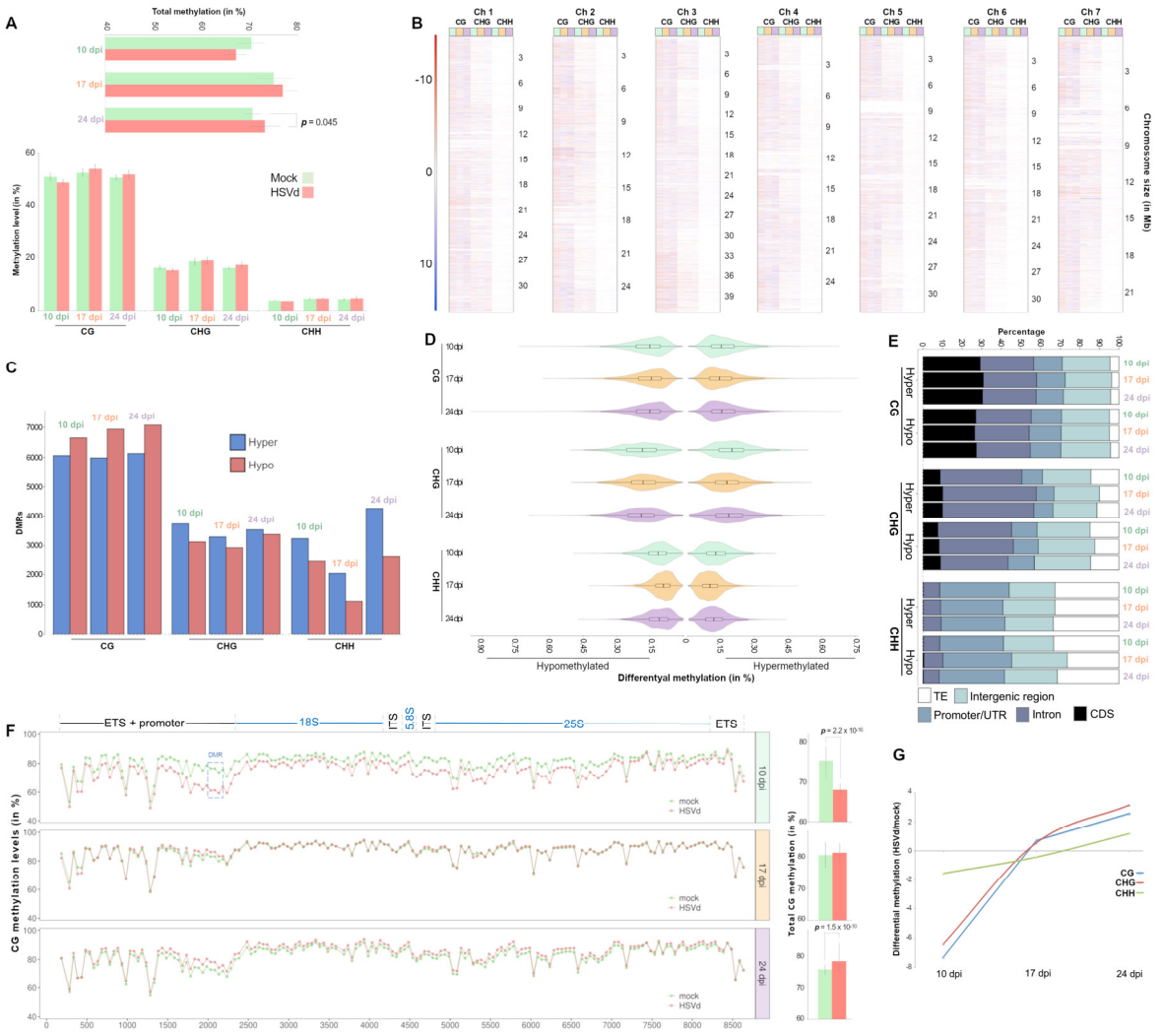


Figure 3.5. HSVd induces alterations in the cucumber epigenetic landscape. A) Graphic representation of the total cytosine methylation mean values of HSVd-infected and control plants at 10, 17 and 24 dpi (upper panel) and discriminated by sequence context (lower panel). Error bars indicate the standard error between replicates. **B**) Chromosome view of the temporal dynamics of the differential methylation profiles in HSVd-infected plants at 10 (green), 17 (orange) and 24 dpi (purple). Red and blue lines represent hypomethylated and hypermethylated regions, respectively. **C**) Number of significant hyper- or hypo- differentially methylated regions (DMRs) in the three sequence contexts (CG, CHG and CHH) identified in infected plants at 10, 17 and 24 dpi. **D**) Violin plot representing the temporal profile of the epigenetic alterations (estimated by significant DMRs) associated to HSVd-infection. Internal box line indicates the median of the percentage of differential methylation values. **E**) Temporal description of the proportion of the cucumber genome regions containing significant DMRs identified in infected plants. **F**) Global view of the methylation profiles in CG context in ribosomal RNA transcriptional unit at 10, 17 and 24 dpi in HSVd-infected (red) and control (green) plants. In the left part each dot represents the mean methylation values of 50 nucleotides while in the right the total mean methylation values of the whole ribosomal unit are represented. A significantly hypomethylated DMR identified in infected plants at 10 dpi and matching to promoter region is highlighted. The ribosomal genes (18S, 5.8S and 25S) are indicated in blue while the internal transcribed spacers (ITS) and external transcribed spacers (ETS) are in black **G**) View of the temporal dynamics of the differential methylation in the three sequence contexts (CG, CHG and CHH) observed in ribosomal RNA transcriptional unit during the HSVd infection.

It has been previously demonstrated that HSVd infection promotes alterations in the epigenetic landscape of regulatory regions of the ribosomal DNA (rDNA) (Martinez et al., 2014) to reconfigure a functional scenario that encourages its replication in cucumber plants (Castellano et al., 2016b). To extend this study, we analyzed if rDNA was specifically identified as a significant

DMR. A region of 150 nt comprised between positions 1984 – 2133 of the ribosomal genes and overlapping with their promoter region was identified as hypomethylated DMR in CG and CHG contexts at 10 dpi (boxed region in Figure 3.5F and Supplementary Figure 3.9). Moreover, considering the total methylation levels at CG context, it was evident that HSVd induces a significant ($p = 2.2x10^{-16}$) hypomethylation of the rDNA (79.57 and 73.75 of total methylation for mock and HSVd-infected plants, respectively) at 10 dpi. Although the downregulation of the methylation levels was more evident in the promoter region (identified as a significant DMR), the hypomethylated status was extended to the entire ribosomal gene (Figure 3.5F – upper panel). Epigenetic alteration of the rDNA associated to HSVd infection was revealed as a dynamic phenomenon, being comparable to mock at 17 dpi (Figure 3.5F – central panel) and significantly ($1.5x10^{-10}$) hypermethylated at 24 dpi with total-CG methylation levels of 80.94% and 83.37% for mock and HSVd-infected plants, respectively (Figure 3.5F – lower panel). This dynamic alteration of the host epigenetic landscape was also observed when CHG and CHH sequence contexts were considered (Figure 3.5G and Supplementary Figure 3.10). Under a temporal viewpoint, the significant hypomethylation of ribosomal genes was coincident with the initial phase in which HSVd reaches a systemic infection, while the transition to an hypermethylation status in host rDNA was in parallel with the stabilization (or slight decrease) of the HSVd transcripts in infected plants (Supplementary Figure 3.11).

Interplay between methylation and transcriptional regulation in infected plants

To analyze if the changes in the methylation levels observed in HSVd-infected plants could be associated to transcriptional alterations we integrated the data obtained by RNA-seq and WGBS assays. 113 protein-coding transcripts with strong (Log2FC ≥ 1 or ≤ -1) differential expression at 24 dpi containing at least one DMR in any of these regions (promoter/UTR, CDS and intron), were selected for this analysis (Supplementary Table 3.10). We observed that 68 of these 113 genes — representing a significant ($p = 0.038$ in Exact binomial test) proportion (60.2%) of the analyzed transcripts— were characterized by containing at least one antagonistic DMR (hypomethylated for over accumulated transcripts or hypermethylated for downregulated transcripts) (Supplementary Table 3.11). Interestingly, a significant negative correlation was observed when we compared their expression levels (estimated by Log2FC) with their global methylation status (determined by DMR analysis in the three sequence contexts) in promoter/UTR and CDS regions ($p = 0.003$ and 0.014, respectively) (Figure 3.6A). No significant correlation was observed for DMRs matching to introns ($p = 0.148$). A more detailed analysis (focused on each context separately) evidenced that the negative correlation was mainly associated to DMRs in the CG context (Figure 3.6B and Supplementary Figure 3.12). Next, we compared the expression and methylation levels of these 68

genes at 10, 17 and 24 dpi to obtain a global overview of the temporal dynamics of the interplay between cytosine methylation and transcriptional response during HSVd-infection. Our results indicate that while the intensity and the density of the global methylation of the analyzed genes (estimated by the mean values and the total number of the DMRs) were increased during the analyzed period (Figure 3.6C, upper panel - left), their transcriptional activity (estimated by the mean of the Log2FC values and the number of DET) was consistently down-regulated (Figure 3.6C, lower panel - left). In a similar way, the transcripts upregulated during HSVd-infection were associated to a temporally increased hypomethylation status of their respective genes (Figure 3.6C, right - lower and upper panel respectively). According to their biological function, the cucumber genes showing methylation levels antagonistic to transcripts accumulation mainly encoded proteins related to membrane (associated to stress response or transport), oxidation-reduction and metabolic processes (Supplementary Table 3.11). Additionally, some genes encoding for nucleus components, transcription factors and protein involved in protein-protein interactions, were also identified.

Regarding the regulation of non-coding transcripts by epigenetic changes, we analyzed the transcriptional activity of the cucumber ribosomal genes, which has been earlier described to be modulated by HSVd infection (Martinez et al., 2014). Our data showed that the significant hypomethylation of rDNA observed at early infection phases was associated to an increased transcriptional activity of rRNA genes determined by the increased accumulation of primary precursors of the ribosomal transcripts (pre-rRNAs) in infected plants (Supplementary Figure 3.13, left panel). During rRNA transcription, RNA polymerase I transcribes length-units pre-rRNAs that are extensively processed into 18S, 5.8S and 25S units by the sequential deletion of external and internal transcribe spacers (ETS and ITS, respectively) (Henras et al., 2008). Consequently, the differential accumulation of pre-rRNAs was estimated by considering the transcripts matching to ITS1 and ITS2 regions. Finally, the over accumulation in infected plants of rRNA-derived sRNAs (rb-sRNAs), an indirect indicator of the transcriptional activation of normally inoperative rRNA transcriptional units (Earley et al., 2010), provided additional support to the link between rDNA hypomethylation and transcriptional deregulation (Supplementary Figure 3.13, right panel).

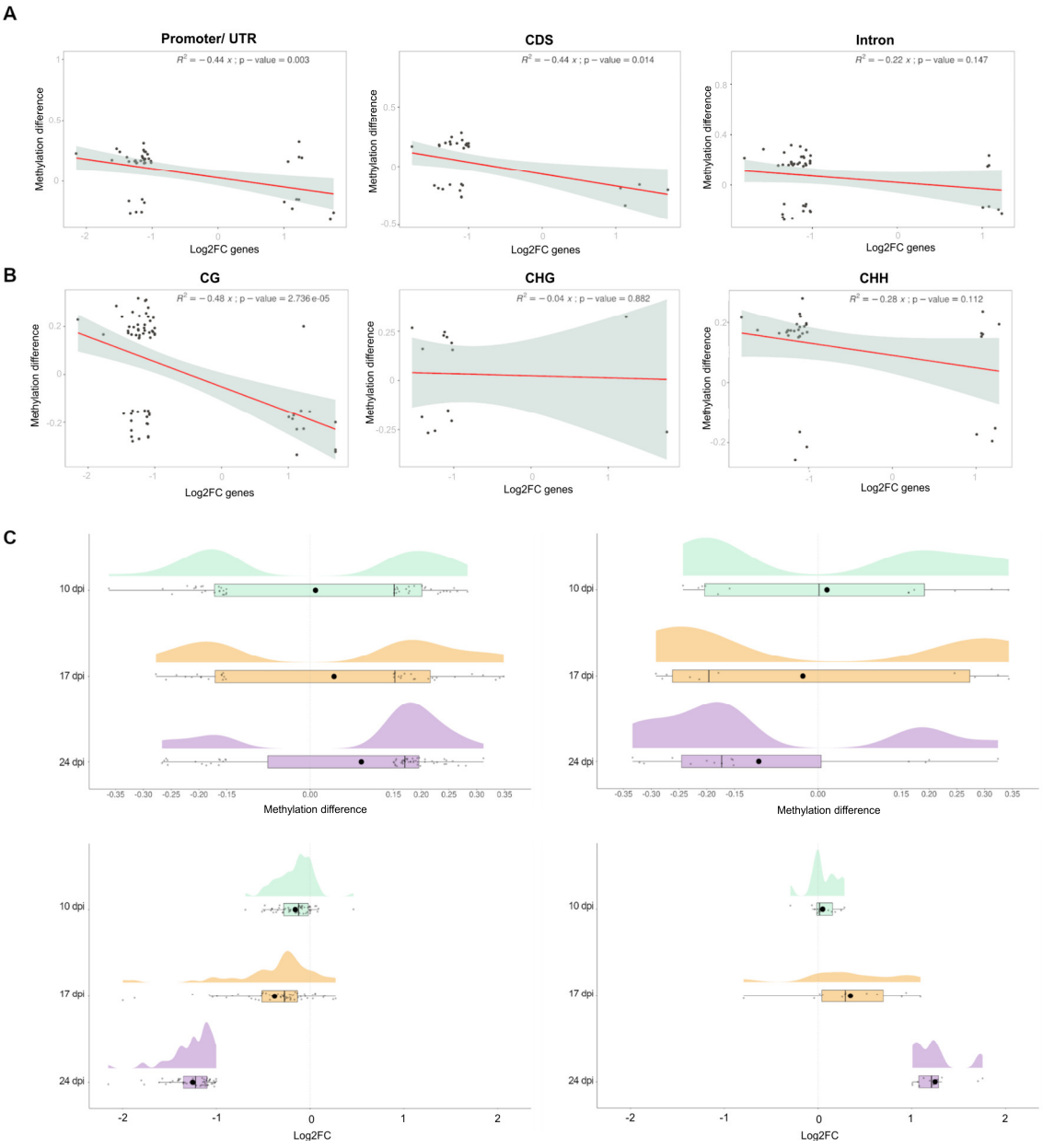


Figure 3.6. Association between transcriptional alterations and epigenetic changes induced by HSVd-infection. Scatter plots showing the negative correlation (significance estimated by Pearson correlation coefficient) between the expression levels of differential genes containing at least one antagonistic DMR and their global methylation status. **A**) Considering the three selected gene regions (Promoter/UTR, CDS and intron) and **B**) considering methylation sequence context. **C**) Raincloud plots showing the temporal correspondence between the sense and the intensity of the epigenetic changes (upper panels) and the transcriptional activity (lower panels) in response to HSVd-infection. The temporal increase in the intensity of the hypermethylated (upper left) and hypomethylated (upper right) DMRs represented by the mean of the differential methylation (black dots) is associated with an evident decrease (lower right) or increase (lower left), respectively (in intensity) of the transcriptional activity estimated by the LFC mean values (black dots).

Discussion

In coincidence with other cellular parasites, viroids are compelled to subvert all sorts of cellular factors and reprogram host gene expression in order to modulate (into their own benefit) complex plant regulatory networks. Assuming this functional scenario, it is expected that the type of affected host-regulatory pathways as well as the intensity of the induced alterations may vary considerably during the infection process. However, our knowledge about the functional diversity and temporal dynamics of the global plant response to viroid infection is limited (Adkar-Purushothama and Perreault, 2020; Navarro et al., 2021a).

Here, we have addressed this question by means of a temporal analysis of the global response to HSVd infection focused on three host regulatory levels small RNA interference, modulation of the transcriptional activity and epigenetic modifications. This integrative analysis allowed us to determine that during the analyzed infection period cucumber plants follow different strategies to modulate their cellular homeostasis in response to the functional alterations induced by HSVd replication and accumulation.

The early detection (at 10 dpi) of viroid-derived transcripts and sRNAs in non-inoculated apical tissues evidenced the capacity of HSVd to develop a relatively quick systemic infection in cucumber plants. However, the observation that viroid transcripts are recovered from only two out the three biological replicates and in a heterogeneous level (27 and 2487 and 2.94 and 1421, normalized reads, for genomic and vd-sRNAs respectively) suggests that the initial infection phase is highly variable and may be dependent on yet unknown viroid-host interactions or that minimal variations on the plant phenological stage could condition these early events.

Analysis of the small RNA data evidenced that, although sRNAs derived from *plus* polarity transcripts were predominantly recovered from infected plants at the three analyzed time points, both *plus* and *minus* HSVd replication intermediates are potential substrates for host RNA silencing machinery at different infection phases. In fact, considering the number of unique sequences no differences were observed between the two polarities. Our results are in consonance with previous studies in which the recovery of small RNAs of both polarities supports the predominant involvement of viroid replication intermediates as source of vd-sRNAs in plants infected by both nuclear (Itaya et al., 2001; Papaefthimiou et al., 2001; Sano and Matsuura, 2004; Navarro et al., 2009; Martinez et al., 2010) and chloroplastic (Martinez de Alba et al., 2002; Di Serio et al., 2009) viroids. Considering the size distribution, 24 nt vd-sRNAs were predominant at the three analyzed time points, followed by 21 nt sequences. However, it was evident that this size-distribution profile was modified during HSVd infection, being characterized by a temporal increasing of the global

proportion of 21 nt that negatively correlates with the accumulation of 24 nt vd-sRNAs. These results support that the processing/accumulation of 24 nt vd-sRNAs is predominantly associated to initial infection steps, while the accumulation of the 21 nt class constitutes an event characteristic of well-established infectious processes. The observation that a similar scenario is not observed when cucumber sRNAs were considered evidenced the independence of this phenomenon with the natural evolution of the endogenous sRNAs biogenesis. This temporal trend may explain the inconsistency between our data and previous results supporting that 21 nt in length is the most abundant class of vd-sRNA in HSVd-infected cucumber plants at late infection phases (Martinez et al., 2010; Zhang et al., 2020). When the distribution of vd-sRNAs along the HSVd-genome was analyzed, we observed that reads matching onto four specific regions represent approximately the 40% of the recovered reads in the three analyzed time points. In contrast, two HSVd-RNA regions exhibited null or extremely lower accumulation of vd-sRNAs. The existence of these hyper- and hypo- accumulating regions (detected in both HSVd strands) may be explained assuming the existence of certain viroid-RNA regions exhibiting differential accessibility to host RNA silencing machinery and/or differential stability rates for these vd-sRNAs. It is interesting to note that one of these hypo-accumulating regions, the terminal left (TL), contains the terminal conserved hairpin (TCH), a highly conserved region in the HSVd sequence (Amari et al., 2001) and that besides *Hostuviroid* is also present in *Cocaviroid* and *Coleviroid* genera. In the *Pospiviroid* genus, this region harbors the DNA-dependent RNA polymerase II binding region and the transcription initiation site (Sano, 2021).

Increasing evidence supports the involvement of vd-sRNAs in the downregulation of host genes via RNA silencing mechanism (Wang et al., 2011; Navarro et al., 2012b; Eamens et al., 2014; Adkar-Purushothama et al., 2015; Avina-Padilla et al., 2015; Thibaut and Claude, 2018). Our data support that HSVd-sRNAs (180+) and (234-) might be incorporated into the host RNA silencing mediated pathway promoting the RISC-mediated degradation of cucumber homologous to RTM2 and AQP, respectively. RTM1 and RTM2 genes encode a jacalin-repeat protein and a transmembrane-containing small heat-shock-like (hsp20-like) protein, respectively, and are involved in a functional mechanism responsible of the specific restriction of the long-distance movement of *Tobacco etch virus* in *A. thaliana* (Kloth and Kormelink, 2020). Although the basis of this phenomenon is unclear, it has been proposed that the restriction could be the result of physical blockage of virus entry into, passage through, or exit from the phloem (Whitham et al., 2000). On the other hand, aquaporins are membrane proteins ubiquitously present in all kingdoms of life. Although were originally discovered as water channels, increasing evidences support their involvement in response to biotic stress (Afzal et al., 2016), probably by regulating the transport of H2O2, a reactive oxygen species essential for pathogens defense, in infected plants (Dynowski

et al., 2008). The precise modes of aquaporin regulation during infection are still unknown (Afzal et al., 2016). However the observation that aquaporin-derived RNAs are differentially recovered from a degradome dataset obtained from tomato plants infected con CEVd (Thibaut and Claude, 2018), suggests that this differential accumulation of aquaporin transcripts might be modulated by alteration in RNA silencing pathways associated to viroid infection. According to the functional characteristics previously proposed for both RTM2 and AQPs, it is feasible to speculate that the vd-sRNA-mediated silencing of their homologous in cucumber might favor the vascular viroid movement (Pallás and Gómez, 2017) and be closely related to interferences in the signaling processes associated to plant defense mechanism. However, further studies are needed in order to validate the reliably of this very exciting idea.

Having established the global landscape of pathogen-derived sRNAs, next we analyzed the influence of the HSVd infection in the endogenous sRNA population. A first view evidenced that the general profile of cucumber sRNAs is not affected in a significant manner in any of the analyzed time points. Considering sequence accumulation specifically, we observed that the biogenesis of endogenous sRNAs remains relatively stable in response to viroid accumulation, with only 1482 unique sRNAs with significant altered expression at 24 dpi. The predominant trend of this response was the upregulation of the host-derived sRNAs. Differentially recovered sRNAs were not identified at 10 dpi, while a considerable proportion of the sRNAs with significant differential expression observed at 17 and 24 dpi were derived from ribosomal transcripts and associated to hypomethylation of rDNA induced by HSVd-infection (discussed in depth below). A similar scenario (characterized by a lower reactivity level) was observed when alteration in miRNA population was analyzed. Seven miRNA families were significantly down-regulated in infected plants, and except for miR319 (responsive also at 17 dpi) the remaining miRNAs were differentially expressed only at 24 dpi. The association between miR319 downregulation and infection was also evident considering the number of family members (12 at 24 dpi) that showed comparable differential expression. However, the changes observed in miR319 level were not associated to antagonist alterations in the accumulation of their regulatory targets. It is well established that TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATIN CELL FACTOR (TCP) mRNA is a highly conserved target of miR319 in diverse plant species (Palatnik et al., 2007) and accumulating evidence has revealed that miR319-regulated TCP (MRTCP) genes participate extensively in plant development by regulating hormone metabolism mediated signaling (Fang et al., 2021). Interestingly, it has been previously reported that tomato plants infected with PSTVd (another member in the *Pospiviroidae* family) exhibit a complex array of changes affecting hormone signaling (Owens et al., 2012). However, the lack of transcriptional alterations in MRTCP genes argues against the possibility that such physiological effects can be associated to HSVd

infection in cucumber. No significant correlation was observed when the targets of the other six differentially expressed miRNAs where analyzed evidencing that, in coincidence with the described in other plant-viroid models (Martín et al., 2007; Tsushima et al., 2015; Zheng et al., 2017; Navarro et al., 2021b), the global effects of HSVd infection in cucumber miRNA metabolisms is weak, and generally not related to an evident regulation of their targets.

Since the first viroid was described, it has been proposed that this type of naked exogenous RNAs may act as an abnormal regulator affecting host transcriptional activity (Diener, 1971). Furthermore, the alteration of the accumulation of plant transcripts, at specific infection phases, has been previously reported in diverse viroid-host interactions (Itaya et al., 2002; Owens et al., 2012; Rizza et al., 2012; Herranz et al., 2013; Zheng et al., 2017). Our temporal analysis of the evolution of the cucumber transcriptional landscape in response to HSVd infection indicates that although significant transcriptional alterations were observed in the three analyzed time points, the intensity of this response (considering both, the number of differential transcripts and the alterations in the accumulation levels) paralleled the development of the infection, being slight at 10 dpi and more evident at 24 dpi. Another particularity associated to the infection time was the trend of the transcriptional response, while the upregulation of host transcripts was the predominant response at 10 dpi, the downregulated transcripts were predominant at 17 and 24 dpi. According to the functionality, it was evident that HSVd-infection is associated to an extensive reprogramming of the host transcriptional activity, affecting cucumber genes involved in diverse cellular functions. This observation is in agreement with previous data obtained from plants infected with two (severe and mild) HSVd-variants, showing a complex array of changes in the host transcriptome associated to viroid infection (Xia et al., 2017). However, a more defined functional response to HSVd-infection was evident when we considered only transcripts differentially expressed at 17 and 24 dpi. Cucumber genes included in this common response were predominantly downregulated and mainly involved in membrane metabolism, transport and response to biotic stress.

Besides transcript accumulation, it is well established that alternative splicing is a regulated process that increases the transcriptome diversity, providing an alternative way to favor the plant adaptation to environmental changes (Shang et al., 2017; Jabre et al., 2019). However, our knowledge about the alternative processing of host transcripts in response to viroid infection is very limited and only data related to PSTVd-infected tomato plants at late infection has been reported (Zheng et al., 2017). Our results evidenced that the differential exon usage was the predominant transcriptional response during the initial phase of the HSVd-infection (10 dpi). In contrast to the observed for transcripts expression, the global trend (up or downregulation) in exon accumulation was comparable at the three analyzed time points. Only a small proportion (5,76%) of the transcripts

having AS, also exhibited significant differential accumulation, suggesting that the control of the transcriptional activity and the alternative processing of pre-mRNAs might be complementary mechanisms activated in cucumber plants in response to HSVd infection. Regarding to the functionality of the transcripts showing significant AS in response to HSVd-infection, we observed that although the initial response included cucumber genes involved in many cellular processes, transcripts showing a more constant differential exons usage frequency during infection were predominantly related to primary metabolism. The observation that the AS of transcripts with similar biological functions was also reported in PSTVd-infected tomato plants (Zheng et al., 2017), suggest that the regulation of the plant metabolism mediated by differential exon usage might be a host response mechanism also extended to other nuclear-replicating viroids.

The differential methylation of cytosines is another host regulatory layer susceptible to be affected in response to viroid infection (Wassenegger and Dalakouras, 2021). Although, this epigenetic mechanism was first described studying viroid infection in PSTVd-expressing transgenic tobacco plants (Wassenegger et al., 1994), the global effects of viroid-infection into host DNA methylation, remain yet unexplored (Navarro et al., 2021a). The WGBS analysis evidenced that HSVd infection is associated with significant alterations in the host epigenetic landscape. Under a global viewpoint, the alteration in cytosine methylation was a dynamic phenomenon associated to infection phases (slight global hypomethylation at early infection and significant hypermethylation at 24 dpi) and sequence context (predominant hypomethylated DMRs at CG context and hypermethylated CHH). Studying in detail the regulatory effects of this altered epigenetic landscape; we observed that the changes in methylation levels (at CG sequence context) of putative promoter regions and CDS correlate with the transcriptional alterations observed in infected plants. Thus, suggesting that these HSVd-associated epigenetic variations may be responsible of the differential expression of certain genes in infected cucumber plants. Previously, it was reported that in cucumber (Martinez et al., 2014; Castellano et al., 2016a) and *N. benthamiana* (Castellano et al., 2015) plants, HSVd accumulation is associated to significant hypomethylation of the promoters region in rRNA genes. The results obtained here, besides reinforcing the previous results, demonstrated that the alterations in rDNA methylation are also extended to the totality of the rRNA transcriptional unity which is in coincidence with the global trend in the whole genome.

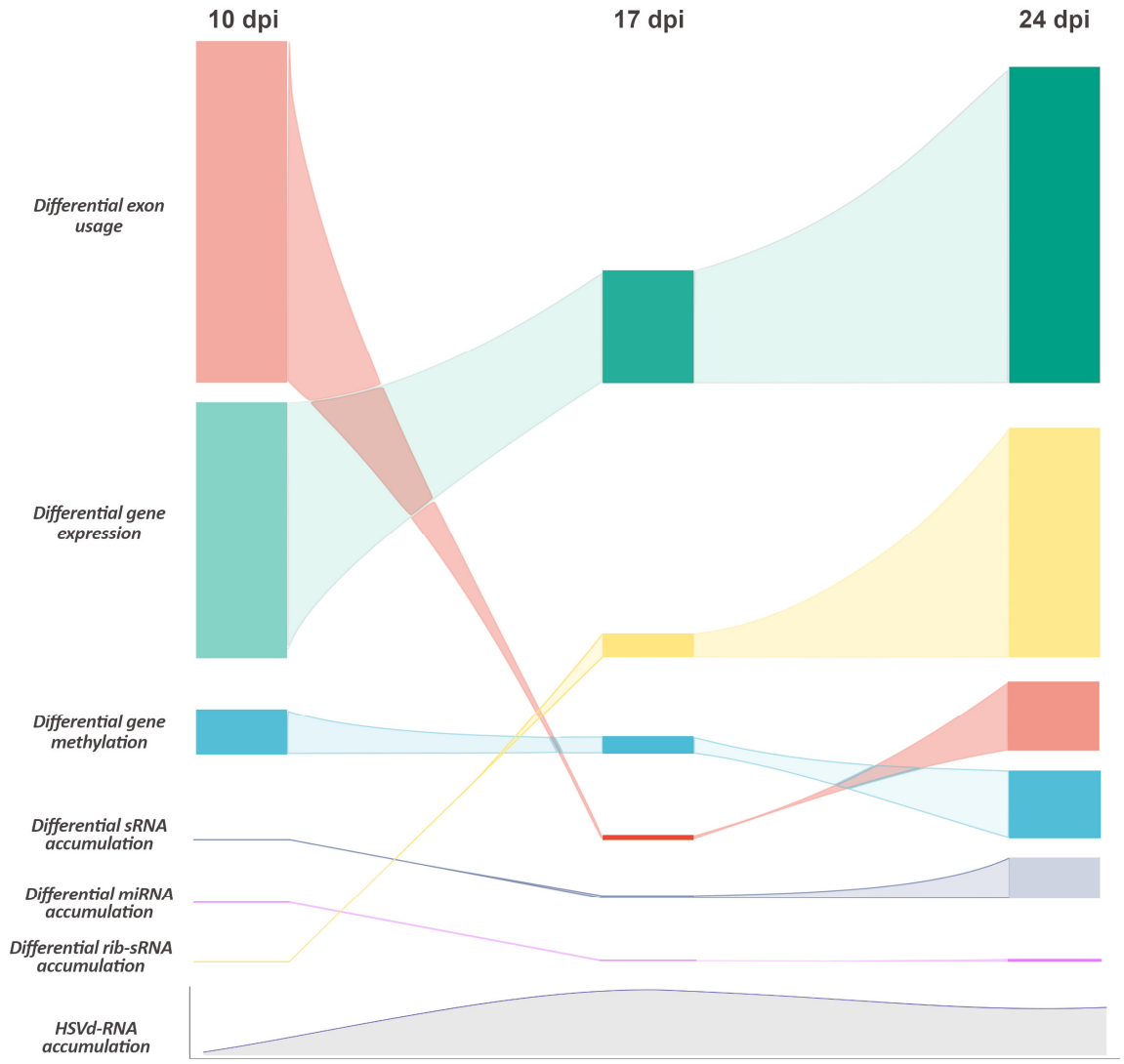


Figure 3.7. Graphic representation showing the temporal dynamics of the global host-regulatory response to HSVd-infection in cucumber plants. The predominant regulatory layers affected in each analyzed infection time are hierarchical represented by the density (box length) estimated by the number of alterations and intensity (box color strength) estimated by the absolute differential value of the response. To represent the density of the alterations in methylation we considered the number of genes with significant DMRs in CG context. In the sRNA group are included all the identified classes except rib-sRNAs and miRNAs. The global position in each analyzed time-point indicates the predominance of the regulatory layer in the global response to infection. Below it is also represented the relative accumulation (estimated by the amount of HSVd transcripts) in infected plants.

The integrative analysis of our data supports that in response to HSVd significant alterations are triggered in cucumber plants at different regulatory levels that are closely related to the temporal dynamics of the infection process (Figure 3.7). Differential exon usage, modulated by alternative splicing, emerges as the predominant response at the initial infection stage. It is well established that the alternative processing of the pre-mRNAs increases the diversity of the plant transcriptome and proteome (Shang et al., 2017). Consequently, this rearrangement of the host-transcriptome (without new transcriptional activity) might, in similar way to the previously described for other

stress conditions (Jabre et al., 2019), allow the immediate fine tuning of the host gene expression in response to HSVd-infection. At subsequent infection phases, the modulation of the plant-response to the viroid will be mainly promoted by alterations in the transcriptional activity that was evidently increased at 17 dpi and being the mostly affected regulatory-layer (considering both, intensity and diversity) at 24 dpi. Interestingly, altered transcription shows a temporal coincidence with changes in the methylation profiles of promoter regions in certain cucumber genes. This functional connection permits to speculate about the possibility that the non-immediate transcriptional response to HSVd might be modulated by the host epigenetic changes associated to infection. Finally, it was obvious that the alteration in sRNA and miRNA metabolism played a minor functional role in the recovering of the cell homeostasis in response to HSVd infection.

An interesting emerging question is how HSVd may trigger the regulatory alterations observed in infected plants. It is generally proposed that disease development must be the result of the direct interaction between viroid-RNA and certain host factors (Adkar-Purushothama and Perreault, 2020). For example, it has been recently shown that the direct interaction between two symptomatic viroids (PSTVd and CEVd) and the host translational machinery induce ribosomal stress in tomato plants (Cottilli et al., 2019). Assuming this possibility, it may be hypothesized that resembling the described for the arabidopsis long ncRNA ASCO (Bardou et al., 2014), HSVd (an exogenous lncRNA) might bind to specific splicing factors affecting pre-mRNA processing. Regarding transcriptional alterations linked to host epigenetic changes, it has been proposed that the direct interaction between viroid-genome and HISTONE DEACETYLASE-6 (HDA6) is responsible of the methylation changes observed in the rDNA of HSVd-infected cucumber plants (Castellano et al., 2016b). Consequently, the possibility that the HSVd might modulate the host epigenetic landscape by sequestering key components of the RNA-directed DNA Methylation (RdDM) pathway should not be yet excluded. Alternatively, the specific silencing of certain host genes mediated by vd-sRNAs can also contribute to the global redesign of the host regulatory structure. Further studies focused on a deep analysis of the potential interactions between viroid-RNA and host components will be needed for understanding the ways employed by HSVd to redesign the cucumber regulatory scenario.

Altogether, the data showed here constitute a comprehensive multiomic approach aimed to provide an overview of the temporal dynamics of the host response to HSVd infection. Recognizing that this is only the initial step in a long road, we expect that this innovative viewing will contribute to elucidate the mechanistic and the molecular basis of the host changes triggered by viroid infection.

Material and methods

Viroid infection and sample collection

Cucumber plants cv. Marketer at cotyledon stage were inoculated with *Hop stunt viroid* (Y09352.1) as previously described (Marquez-Molins et al., 2019). Both cotyledons were agro-infiltrated with a culture of *Agrobacterium tumefaciens* strain C58 with 1 of optical density, harbouring pMD201t-HSVd or empty vector, diluted in infiltration buffer (MES 0.1 M, MgCl2 0.1 M). Plants were kept in a photoperiod of 16 h under visible light and 30 °C (light)/25 °C (darkness). The second leaf after the apex was collected at 10-, 17- and 24-days post inoculation. Each bio-replicate consisted of the leaves from three plants and for each time-point three bio-replicates were sampled of mock and HSVd infected plants.

DNA and RNA extraction and library preparation.

Total RNA was extracted using TRIzol reagent (Life Technologies) following the manufacturer instructions. Genomic DNA was extracted using a CTAB-based extraction method as previously described (Healey et al., 2014). All libraries were constructed by Novogene Europe (Cambridge, U.K.) according to their standard procedures, including the previous steps of mRNA purification with oligo(dT) beads sRNA purification by size selection.

Small RNA sequencing and analysis

sRNA libraries were sequenced by single end (50 pb). The resulting reads were adapter trimmed and filtered on quality by Cutadapt v2.8 (Martin, 2011) and Trimmomatic v0.32 (Bolger et al., 2014), respectively. Additionally, only reads in the 20-25 length range and without indeterminations were kept using a custom script. To study the correlation exhibited by the sRNA expression profiles among the different samples, principal component analysis (PCA) was used. PCA was performed using the prcomp function with scaling in the stats R package v. 4.0.4 (R Core Team 2013). Mann-Whitney-Wilcoxon tests were performed to assess for significant differences in the data clusters for Euclidean distances calculated between groups and among groups with the wilcox.test function in the stats R package. Libraries were aligned to the genomes of HSVd (Y09352.1) and *Cucumis sativus* Chinese Long V3 (Zheng et al., 2019b) to determine the provenance of every sRNA. Only sequences that aligned with no mismatches exclusively either to the viroid or the plant genome with a minimum of 0.5 RPMs were considered for further analysis. Additionally, for estimating the size abundance, only samples with all the sRNA lengths were considered. Moreover, HSVd-sRNAs were classified according to their origin (positive or negative viroid strand) considering unique sequences and also the accumulation (rpm). Then, an exact binomial test was performed to assess the significance. Differential expression

of cucumber sRNAs was estimated using two R packages DESeq2 (Love et al., 2014) and edgeR (Robinson et al., 2010) for pairwise differential expression analysis of expression data. Differentially expressed sRNAs were filtered out using two criteria: (i) adjusted p-value ≤ 0.05 and (ii) RPMs ≥ 5 for at least two libraries in stressed or control samples. P-values were adjusted by False Discovery Rate (FDR). miRNAs were annotated from previously described miRNAs in cucumber (Xu et al., 2020) and validated with the precursors deposited in RNA central (Sweeney et al., 2019). The rest of sRNA categories were annotated using RNA central and sRNAanno (Chen et al., 2021).

mRNA sequencing and analysis

RNA sequencing paired reads (150 bp) were aligned to the cucumber genome using STAR (Dobin et al., 2013). For non-annotated transcripts, a *de novo* assembly was performed using Trinity (Grabherr et al., 2011). 15589 transcripts non-overlapping with the gene annotation of Chinese Long V3 (Zheng et al., 2019b) were aligned with gmap allowing chimeric alignments (Wu and Watanabe, 2005) and a custom annotation of non-coding transcripts was obtained using bedtools (Quinlan and Hall, 2010). HTSeq-counts (Anders et al., 2015) was firstly used to count reads per annotated genes using 10 as a minimum alignment quality and secondly of the novel annotated transcripts with a minimum alignment quality of 0 and counting all reads ambiguously mapped in both categories and wherever they overlap (in order to account for potential transposable elements). The count tables were used in DESeq2 (Love et al., 2014) to infer significant expression, considering an adjusted p-value under 0.05 for significance. Clustering analysis of the transcripts with significant differential expression in the three time-points was calculated with stats R-package. Firstly, the optimal number of clusters was determined using the complete method in the NbClust R-package considering Euclidean distances (Charrad et al., 2014). Secondly, the function '*kmeans'* of stats was applied with one million iterations. The resulting clusters (grouping transcripts with similar evolution of the differential expression profiles over the infection) were generated by ggplot2 (Wickham, 2011). Moreover, to account for differential exon usage, the package DEXSeq was employed (Anders et al., 2012), also considering an adjusted p-value under 0.05 for significance. Volcano plots were created using ggplot2 (Wickham, 2011). GO term enrichment analysis was performed using the tool from Cucurbit Genomics Database (CuGenDB) (Zheng et al., 2019b). p-values were corrected using false discovery rate (FDR) and the cutoff p-value for significant represented terms was 0.05.

5' PARE analysis

Libraries for 5' parallel analysis of RNA end (PARE) were constructed following the protocol previously described with minor modifications (Zhai et al., 2014). Poly (A+) RNA (≈1000 ng) was purified from total RNA from each mock and HSVd replicate of 24 dpi using the GenElute Direct mRNA Miniprep Kit (Sigma-Aldrich). A 5′ RNA oligonucleotide adaptor (PARE 5′ Adapter) containing a Mme I recognition site was ligated to the 5′-phosphate of the truncated poly (A+) RNA by T4 RNA ligase. The ligated products were purified by ethanol precipitation and subjected to a reverse transcription reaction with an oligo dT primer with a known tail (dT primer). Following RNA degradation by alkaline lysis, the cDNA was amplified by PCR using adaptor and oligo dT-specific primers (I5(1)-F and GR3′, respectively), digested with Mme I and ligated to a 3′ double DNA adaptor (dsDNA-top and dsDNA-bottom). The ligated products were separated by electrophoresis in polyacrylamide, and those with the expected size were gel-purified and amplified with I5(2)-F and the corresponding indexed I7-R. PCR products of the expected size were gel-purified and subjected to sequencing by Illumina technology. Primers used are listed in Supplementary Table 3.1. The resulting reads were adapter trimmed by Cutadapt v2.8 (Martin, 2011) and filtered on quality and length by fastp v0.23.1 (Chen et al., 2018b), including only reads in the 19-21 length range. The sequence quality was checked by FastQC v0.11.5. Then, libraries were aligned on the assembled transcriptome of *Cucumis sativus* Chinese Long V3 (Zheng et al., 2019b). Degradome sequences were filtered according to the following: i) alignment on a differentially downregulated transcript at 24 dpi, ii) presence in at least two HSVd libraries and absence in mock libraries and iii) match with the predicted target site of a HSVd-sRNA according to the psRNATarget prediction (Dai et al., 2018).

Reporter assay and RT-qPCR

The predicted sequences to be targeted by HSVd-sRNAs were cloned upstream of the yellow fluorescent protein (YFP) by self-hybridizing oligonucleotides to obtain the reporters REP-CsaV3_1G041760 and REP-CsaV3_6G040100 (listed in Supplementary Table 3.1). These oligonucleotides comprise the start codon ATG followed by eight codons (in frame with the YFP) of the cucumber transcript which include the 21 nucleotides predicted by psRNATarget (Dai et al., 2018). *Nicotiana benthamiana* leaves were agro-infiltrated with a mixture of the reporter with either the binary plasmid expressing HSVd (Marquez-Molins et al., 2019) or the empty plasmid (mock) at optical densities 600 nm of 0.05 and 0.5, respectively. Fluorescence was observed at three days post-inoculation with Leica MSV269 stereomicroscope with an excitation wavelength of 485 nm and images were taken with Leica DFC7000T with identical parameters focusing the infiltration circle. RNA was extracted from the infiltrated leaves after fluorescence observation

using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). First-strand cDNA was synthesized by reverse transcription using oligo(dT)18 and RevertAid cDNA Synthesis Kit (Thermo Scientific™, Waltham, MA, USA). qRT-PCR assays were performed using PyroTaq EvaGreen mix Plus (ROX) (CulteK Molecular Bioline) according to the manufacturer's instructions. All analyses were performed in triplicate on an ABI 7500 Fast-Real Time qPCR instrument (Applied Biosystems). Relative YFP expression (YFP-F, YFP-R) was determined by using the comparative ΔΔCT method and normalized to the geometric mean of F-box (HK-FBOX-F, HK-FBOX-R). The statistical significance of the observed differences was evaluated by the paired t-Test.

Bisulfite sequencing analysis and DMR identification

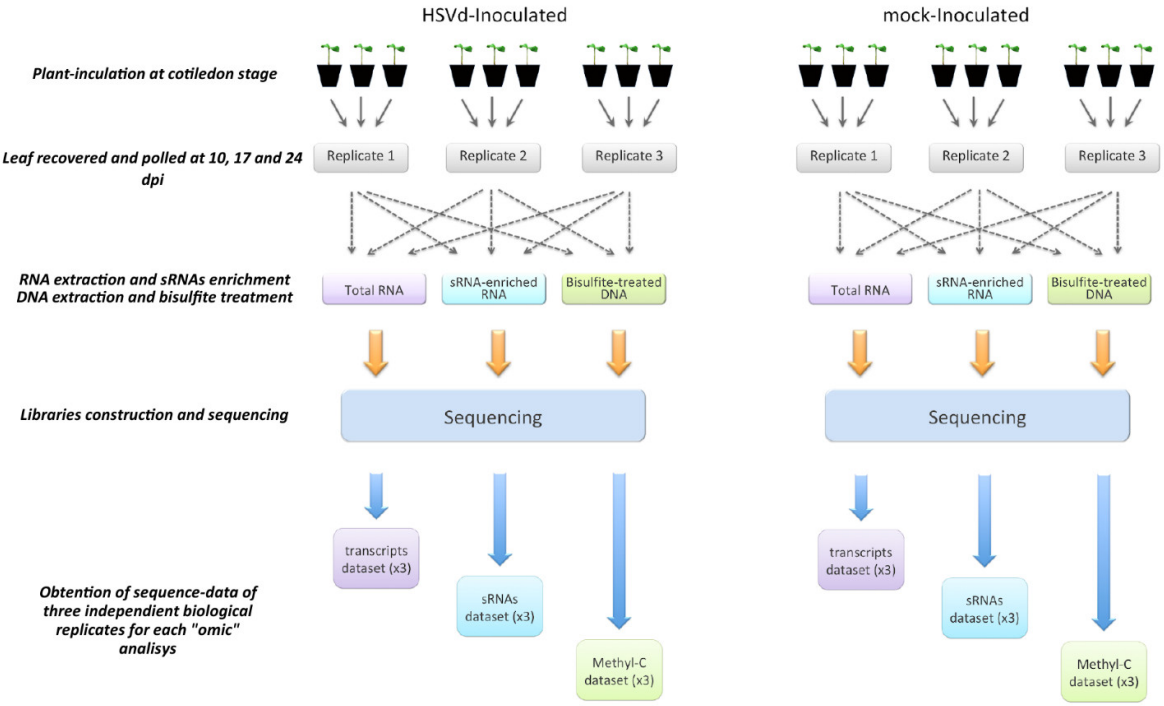
Whole Genome Bisulfite Sequencing (WGBS) libraries were sequenced by paired end (150 pb). The resulting reads were adapter trimmed and filtered on quality by Trimgalore v0.6.6 . Additionally, 10 bases from 5' ends were trimmed from reads using the same software. The sequence quality was checked by FastQC v0.11.9. The mean conversion rate for the 36 libraries was above 99%. Clean reads were mapped to the reference genome using bismark (Krueger and Andrews, 2011) allowing one mismatch per 25 nucleotides length seed. Alignments at the same position were removed using deduplicate_bismark script, including forward and reverse reads. The bismark_methylation_extractor script was used to extract the methylation call for every single cytosine analyzed and obtain a genome-wide methylation report discriminating by context (CG/CHG/CHH).

The analysis of differentially methylated regions (DMR) was carried out with the R-package DMRcaller (Catoni et al., 2018) dividing the genome in equal bins of 50 pb and pooling the samples of the same condition. The DMR were then computed by performing Fisher's exact test between the number of methylated reads and the total number of reads in both conditions for each bin. The obtained p-values were then adjusted for multiple testing using Benjamini-Hochberg's method to control the false discovery (Benjamini and Hochberg, 1995). The criteria chosen to consider a bin as a DMR were the following: i) adjusted p-value <= 0.05, ii) at least three cytosines in the specified context, iii) more than a 15% methylation difference between the two conditions and iiii) at least an average number of reads of eight. Finally bins that were at less than 300 pb were joined.

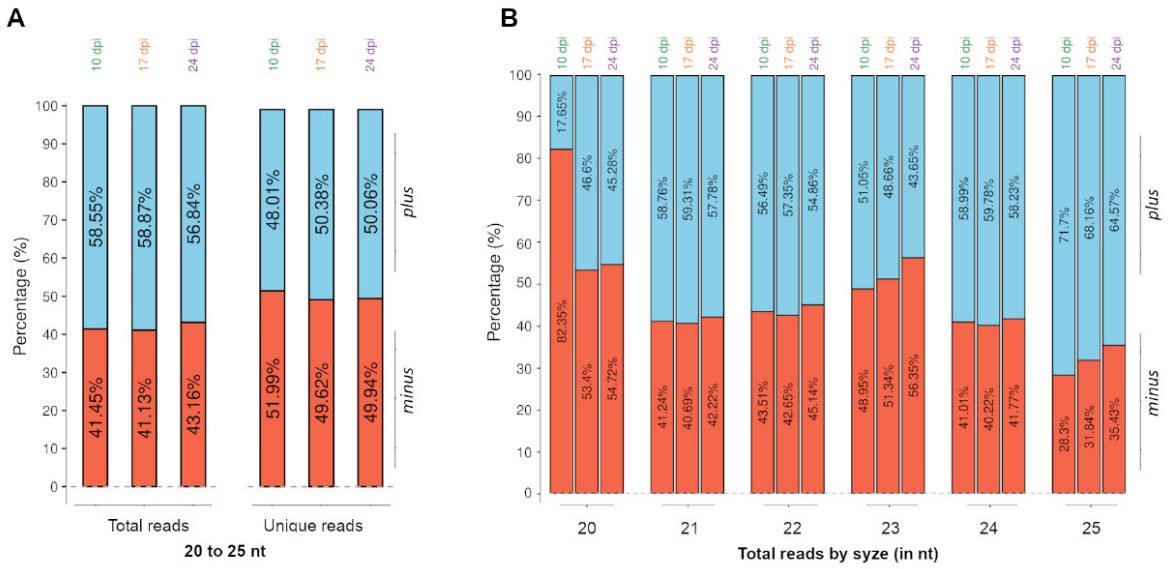
In addition to the DMRs identification, overall methylation was calculated as the mean of the percentage methylation of cytosines distinguishing among contexts. Those cytosines covered by less than eight reads were discarded. Next, a heat-map representing the genome-wide methylation difference between the two conditions was performed. First, cytosine methylation percentages

were grouped in 3000 nucleotide size windows. Second, the methylation difference was calculated according to time and context. Third, the methylation difference of windows meeting the following terms was set to zero: i) mock or HSVd show zero reads coverage and ii) mock and HSVd show less than eight reads coverage. Methylation difference values less than -15% or greater than 15% were considered as -15% and 15%, respectively. Red and blue colors represented the hypomethylation and hypermethylation in HSVd. Both approaches were also carried out for the ribosomal sequence.

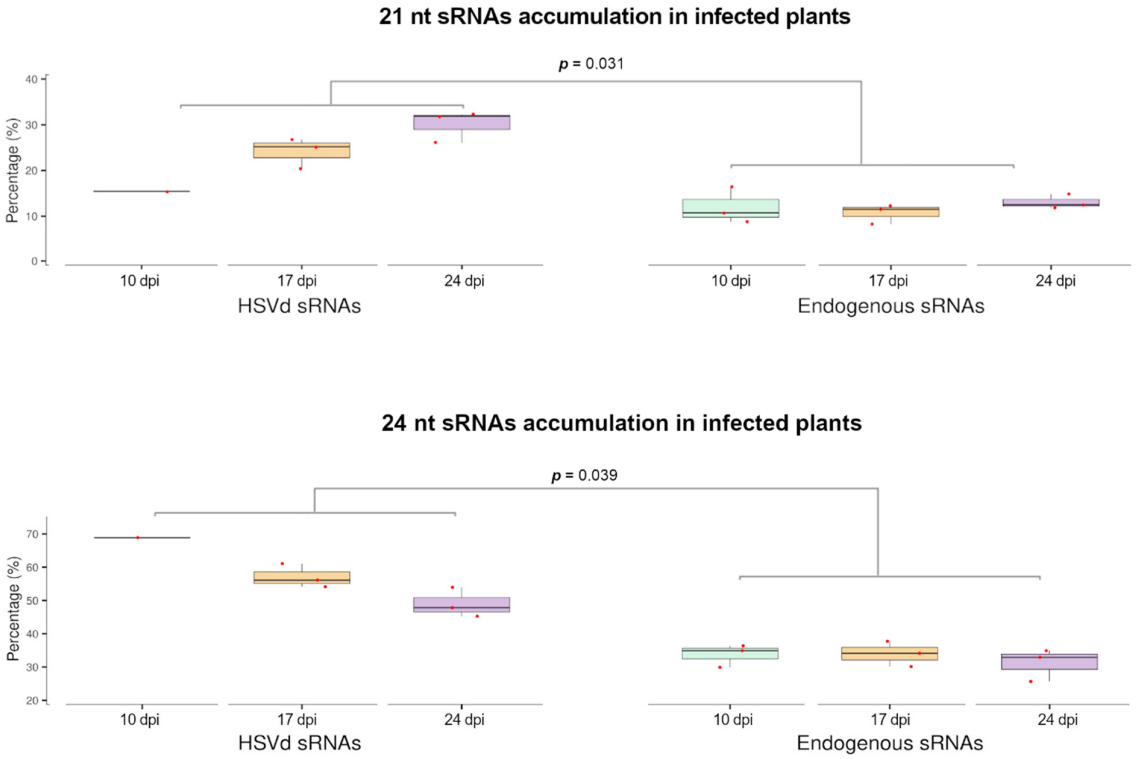
Supplementary material



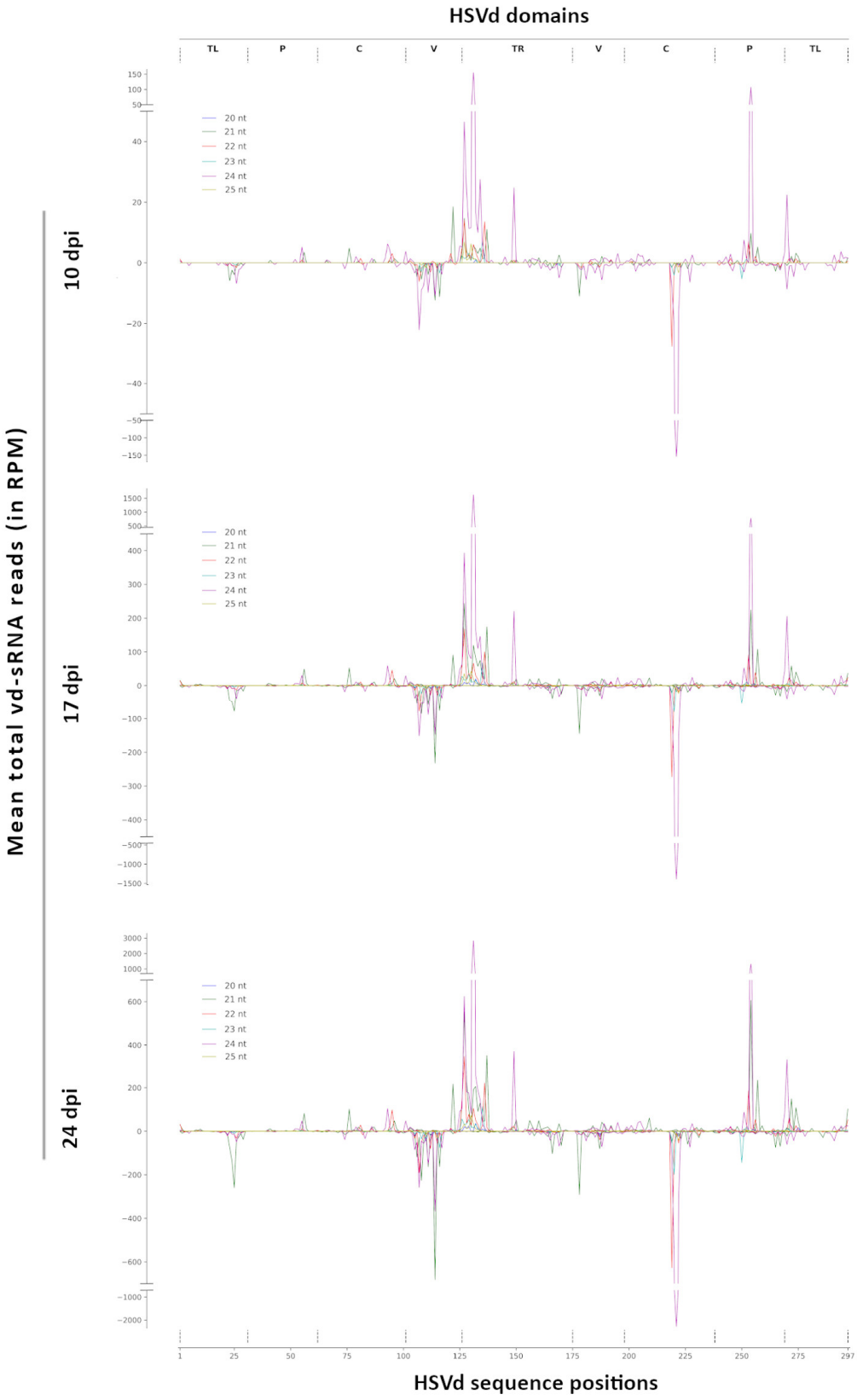
Supplementary Figure 3.1. Graphic representation of the experimental approach used in this study to identify HSVd-induced alterations in the cucumber transcriptome, sRNAnome and methylome. In total 54 independent libraries corresponding to three biological replicates for HSVd- and mock-inoculated samples at three different time points (10 dpi, 17 dpi and 24 dpi) were analyzed.



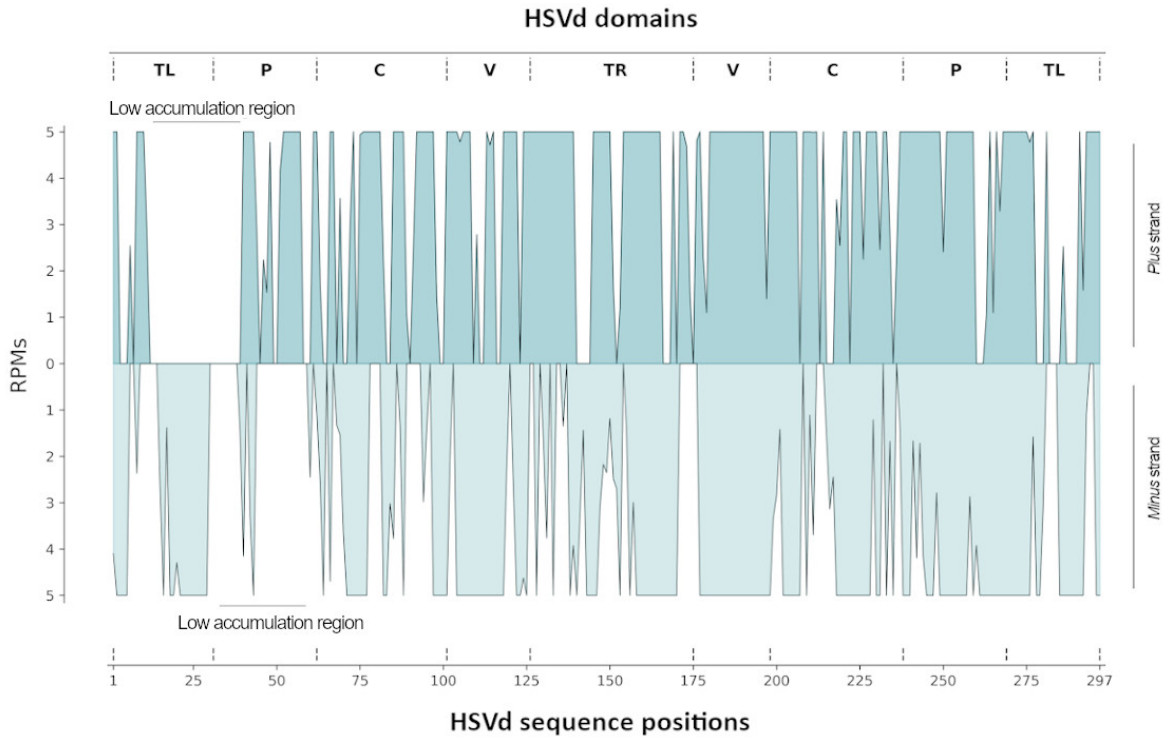
Supplementary Figure 3.2. Viroid-derived sRNAs arise from both HSVd-RNA strands. A) Relative accumulation of the total (left) and unique (right) vd-sRNAs (20 to 25 nt) derived from genomic (plus) and antigenomic (minus) HSVd-RNA strands at the three analyzed time points. **B**) Polarity distribution of the total vd-sRNAs discriminated by infection time and size.



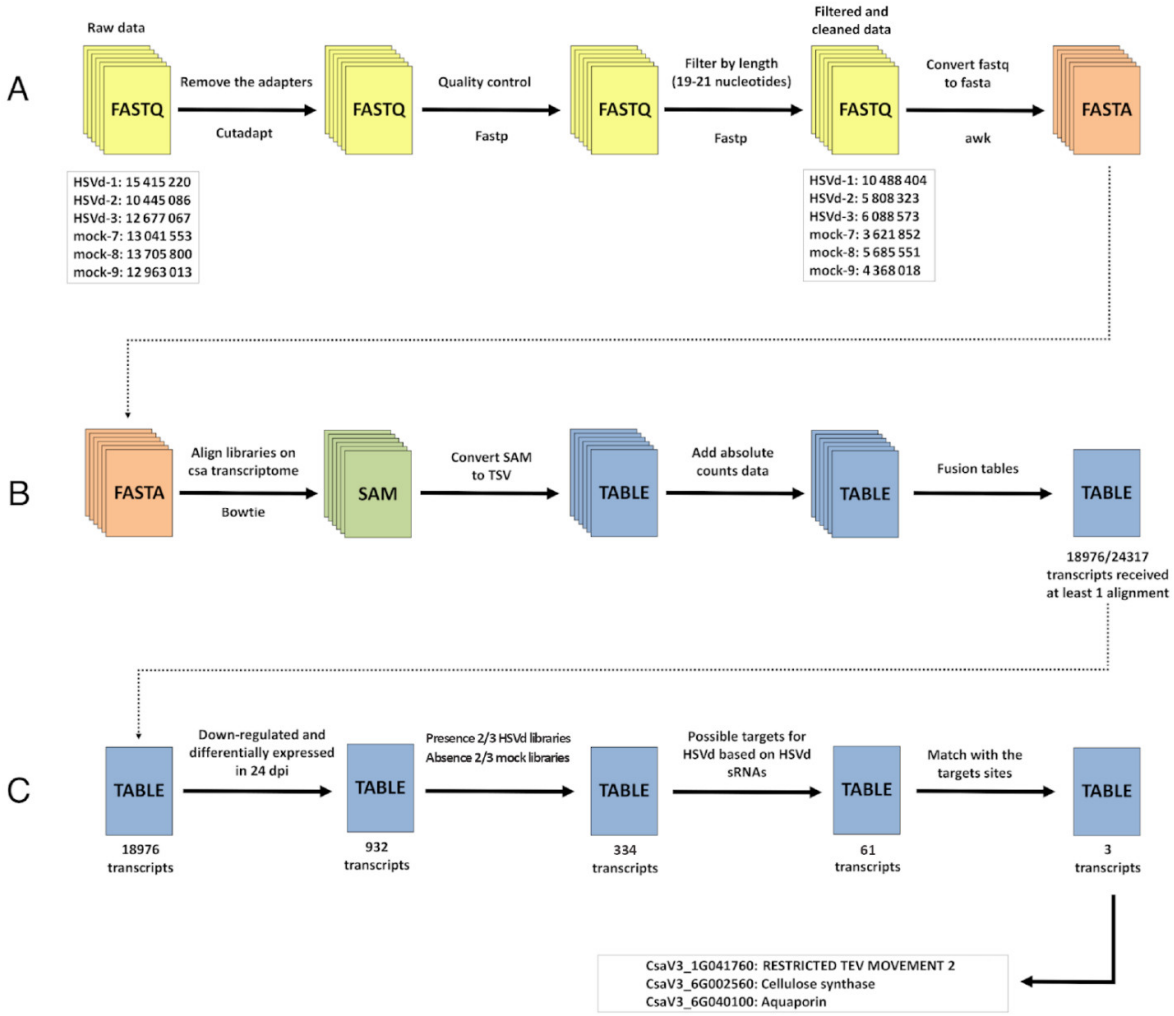
Supplementary Figure 3.3. The relative accumulation of 21 and 24 nt vd-sRNAs evolves differentially during the infection. Box-plot analysis showing a comparison of the mean values (in percentage) observed for 21 nt (upper panel) and 24 nt (lower panel) vd-sRNAs with endogenous sRNAs of the corresponding size at 10, 17 and 24 dpi. The expression of each sRNA type is represented by the median (internal box line) of the percentage of RPM values.



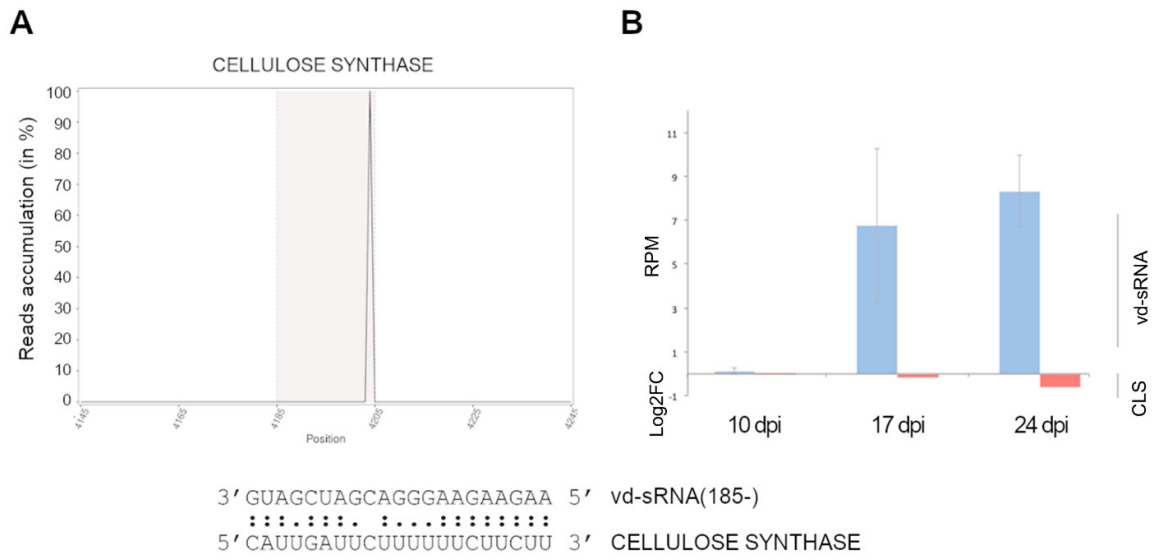
Supplementary Figure 3.4. Genome view of the vd-sRNAs recovered from infected cucumber plants. The vd-sRNAs were plotted according to the position of their 5′-end onto the HSVd- RNA sequence in either sense (above the x-axis) or antisense (below the x-axis) configuration. The values on the y-axis represent the mean of total vd-sRNAs (sizes indicated by colors) in the three analyzed time points.



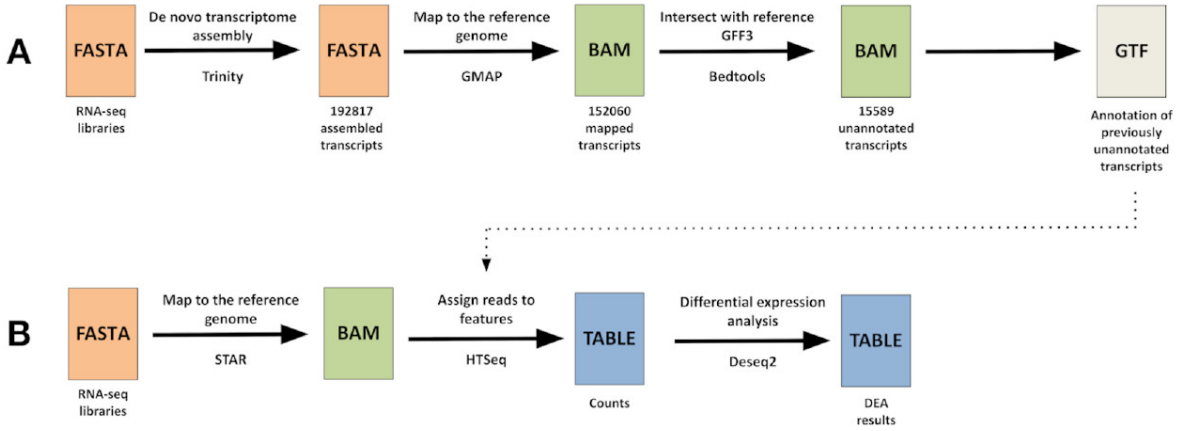
Supplementary Figure 3.5. Two HSVd regions show extremely low accumulation of vd-sRNAs. Detailed genome view of the vd-sRNAs recovered from infected cucumber plants. The vd-sRNAs (21 to 25 nt) were plotted according to the position of their 5′-end onto the HSVd- RNA sequence in either sense (above the x-axis) or antisense (below the x-axis) configuration. The values on the y-axis represent the abundance of total vd-sRNAs limited to 5 RPM in order to evaluate which regions of the HSVd genome contribute to the production of vd-sRNAs.



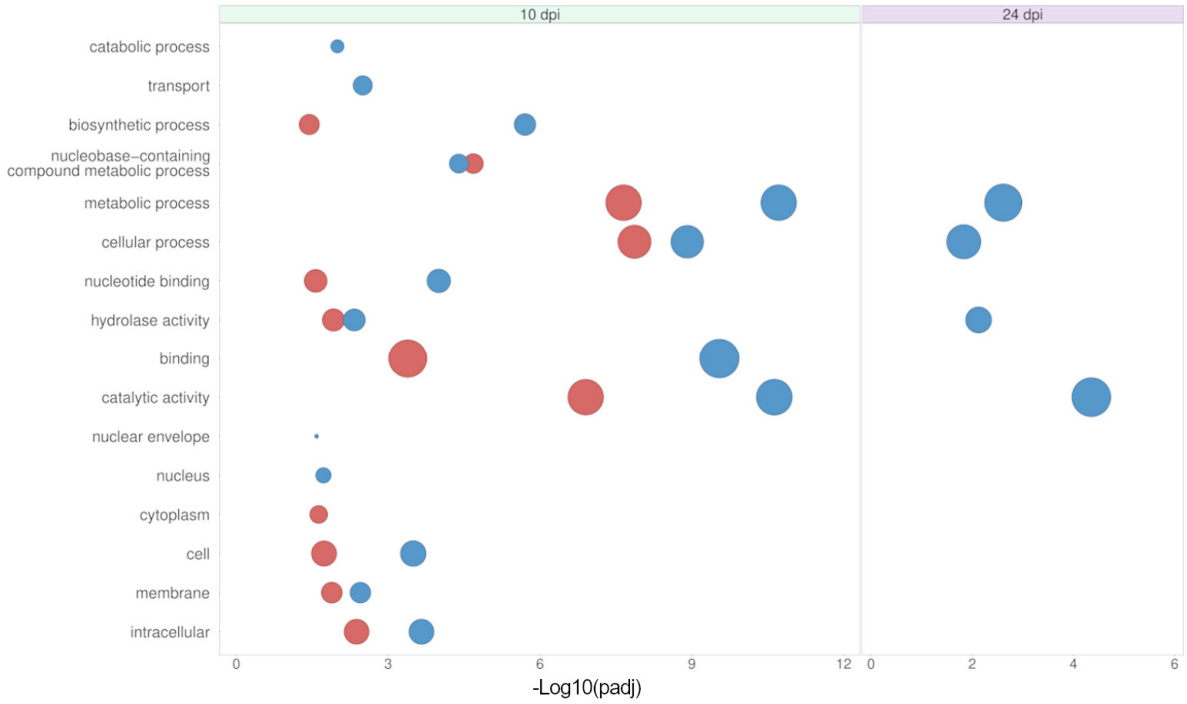
Supplementary Figure 3.6. Details of the pipeline used for the degradome analysis to identify sequences indicative of vd-sRNA-mediated cleavage of cucumber transcripts. A) Filtering of reads and quality control. **B**) Alignment to the cucumber transcripts. **C**) Identification of potential target transcripts.



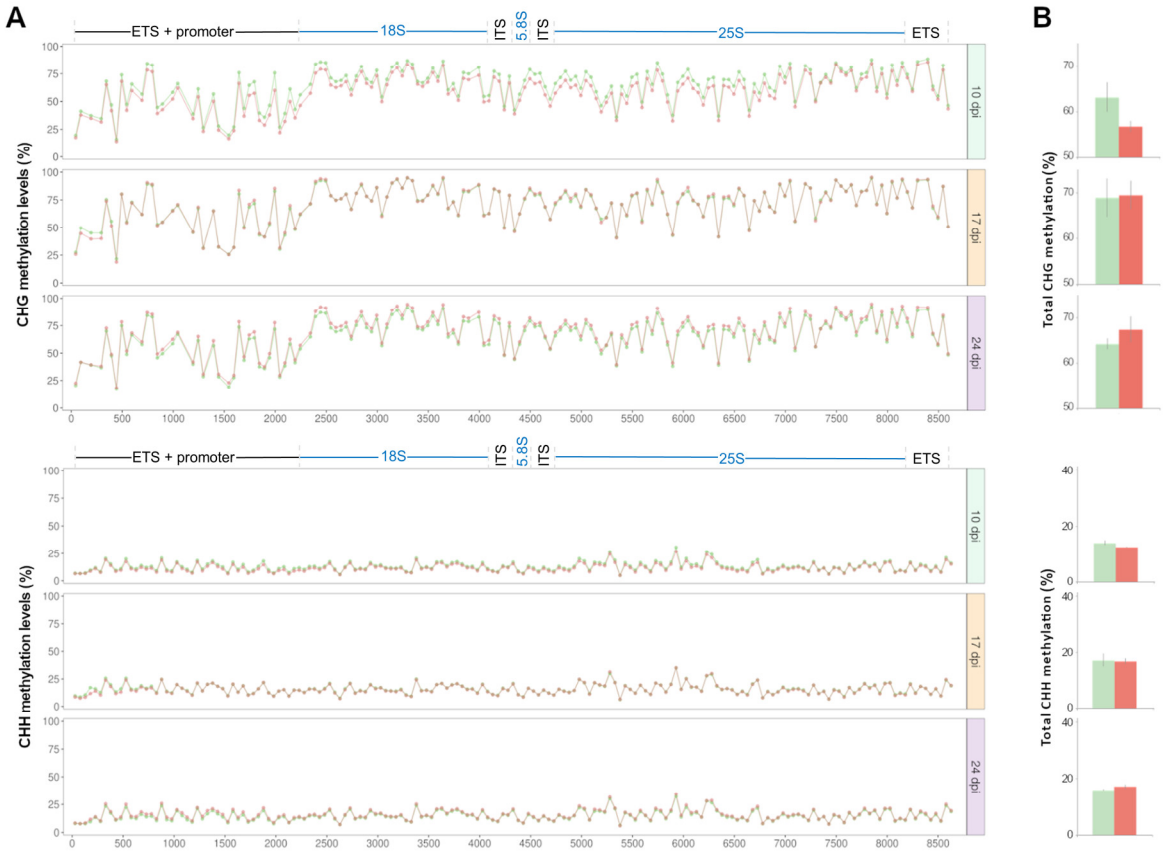
Supplementary Figure 3.7. Representation of the cellulose synthase transcript differentially detected to be cleaved by a vd-sRNA by degradome assay in infected plants. A) Degradome sequences were plotted onto the predicted target transcript. The grey zone indicates the predicted vd-sRNA recognition site (detailed below). The values on the Y-axis represent the relative accumulation of the reads matching onto the 100 nt in length analyzed region. **B**) Dynamics of the accumulation of the vd-sRNA (blue) and its predicted target (red) over the infection. Vd-sRNA accumulation is expressed as the mean of the reads per million (RPM) and the error bars indicate the standard error between biological replicates. The accumulation of the cellulose synthase (CLS) transcript is represented as the Log2FC value.



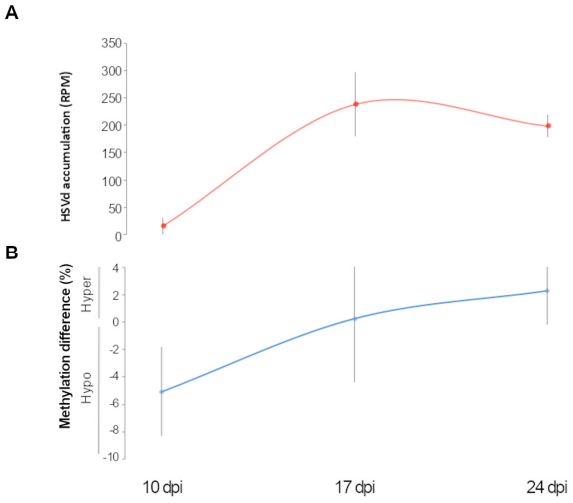
Supplementary Figure 3.8. Details of the pipeline used to identify non-annotated differentially expressed cucumber transcripts. A) Assembly and annotation of previously non-identified transcripts. **B**) Alignment of reads to cucumber genome and differential expression analysis.



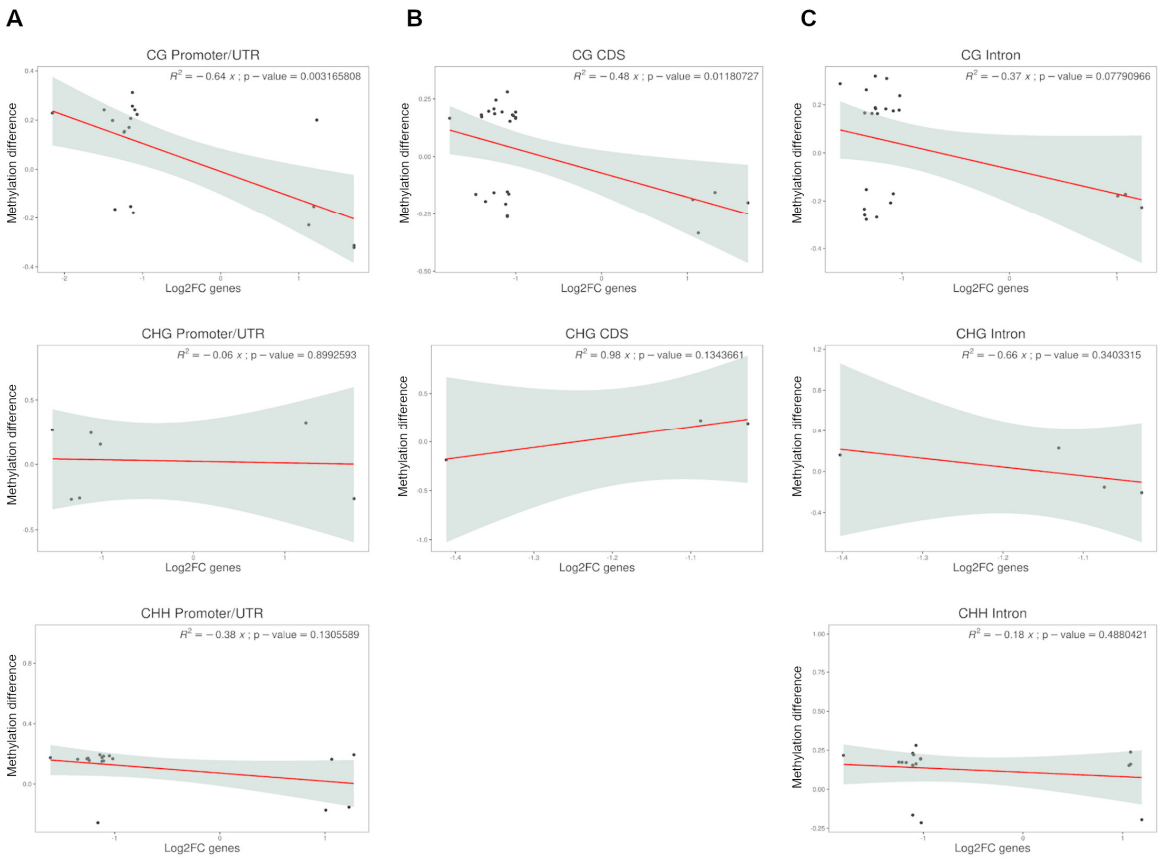
Supplementary Figure 3.9. Gene ontology analysis (plant GOSlim) for transcripts with differential exon usage identified in HSVd-infected plants at 10, 17 and 24 dpi. Circle size is proportional to the fraction of genes relative to the total number of genes with the GO term. Color indicates whether the genes are upregulated (blue) or downregulated (red). The -Log(10) of the adjusted P values is represented in the X-axis.



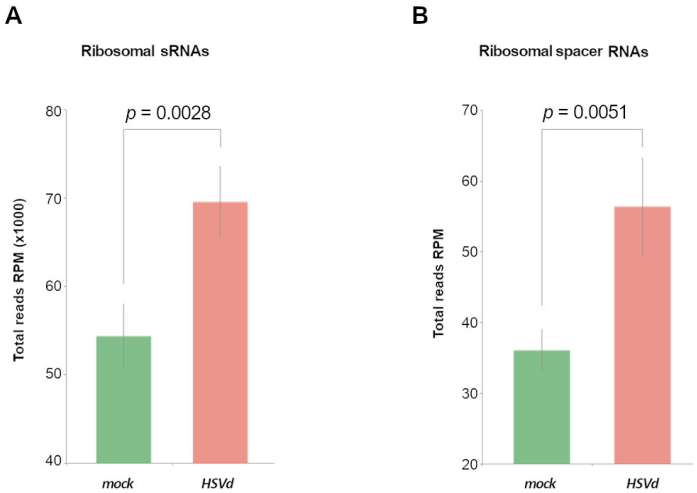
Supplementary Figure 3.10. Global view of the temporal evolution of the differential methylation in ribosomal genes in CHG and CHH contexts. A) Methylation profiles in CHG (upper) and CHH (lower) context associated to HSVd-infection. Red and green dots represent the means of methylation (estimated for 50 nt regions) of rDNA in HSVd-infected and control plants respectively. **B**) Graphic representation of the means of the total cytosine methylation values at CHG (upper) and CHH (lower) sequence context of ribosomal genes in HSVd-infected and control plants at 10, 17 and 24 dpi. The ribosomal genes (18S, 5.8S and 25S) are indicated in blue while the internal transcribed spacers (ITS) and external transcribed spacers (ETS) are in black



Supplementary Figure 3.11. Association between HSVd accumulation and rDNA methylation. A) Representation of the temporal evolution of viroid accumulation estimated by the mean of HSVd-transcript reads recovered in each analyzed time (in red). **B**) Graphic showing the mean differential methylation in infected plants of ribosomal genes at the three analyzed time points. Error bars indicate the standard error between replicates.



Supplementary Figure 3.12. Association between transcriptional alterations and epigenetic changes induced by HSVd-infection. Scatter plots showing the correlation (with significance estimated by Pearson correlation coefficient) between the expression levels of differential genes containing at least one antagonistic DMR and their global methylation considering the three selected gene regions (**A**) Promoter/UTR, (**B**) CDS and (**C**) intron in the three sequence contexts: CG (upper), CHG (middle) and CHH (lower).



Supplementary Figure 3.13. HSVd-infection is associated to significant transcriptional alterations of rRNA. Diagrams showing the mean accumulation of the total ribosomal derived sRNAs (**A**) and transcripts of the ribosomal internal spacers (**B**) recovered from mock-inoculated (green) and HSVd-infected (red) plants during the analyzed infection period. Error bars indicate the SE between biological replicates. Significance values were estimated by paired T-test.

The supplementary tables can be accessed in the following link:

https://www.dropbox.com/sh/lzrmy2gy5dshiu5/AADKc0V7A-TIWYfZ2_Ne3Ycca?dl=0

GENERAL DISCUSSION

A deeper understanding of the nature and mode of action of viroids has been the encompassing main goal of this PhD. For this purpose, simple and efficient procedures for obtaining infectious cDNA clones are essential. A new efficient method for constructing infectious viroid clones was developed and tested with one viroid of each family: ELVd (*Avsunviroidae*) and HSVd (*Pospiviroidae*). This procedure was based on type IIS restrictions enzymes that cut outside of the recognition site. Therefore, it is possible to freely choose the nucleotides left overhang after the digestion with such enzymes, a fact that enables the seamless ligation of DNA fragments in an oriented manner, a strategy that is the base of Golden Gate cloning (Engler et al., 2009). We validated the universality of this approach by calculating the proportion of viroid sequences in which both restriction sites (BsaI and BsmBI) are present, which is indeed extremely low (0.2%). Therefore, this strategy is ideal for cloning oriented dimers of these tiny pathogens. Nonetheless, oriented assemblies of DNA sequences for other purposes could be obtained following similar approaches, especially since repetitive sequences are often difficult or impossible to synthesize. For example, constructs with a short tandem target mimic (STTM) are used to inhibit the action of miRNAs (Yan et al., 2012). An STTM is composed of two short identical sequences mimicking small RNA target sites separated by a 48-nucleotide linker, and thus, since the spacer sequence is always the same, type IIS enzymes could be used to create compatible ends in this spacer to be joined with the target mimicking repeats. Even more, this STTM unit could be dimerized, as we did for viroid sequences, which could potentially inhibit to a greater extent the miRNA but alternatively, STTM targeting different miRNAs could be combined, two issues that have not been addressed yet. Another potential application could be the obtention of constructs with inverted repeats for generating hairpin RNAs that can trigger the methylation of a determined locus, as it has been used in several studies (Dalakouras et al., 2009; Williams et al., 2015; Pignatta et al., 2018). In short, type IIS-based cloning designs are a valuable tool not only for obtaining dimers of viroid sequences but also for many other applications in molecular biology.

Despite viroids have been considered as plant-pathogenic non-coding RNAs since their discovery, our computational analysis predicted three small open reading frames in each of the HSVd and ELVd genomes. No significant similarities with proteins in the database of higher plants were found, but some of these predicted peptides were highly conserved among all HSVd and ELVd variants. For instance, ELVd E-ORF1 showed an identity mean of 81.17% in 96.6% of the accessions whereas H-ORF3 was detected in 59% of the analyzed sequences with an identity value mean of 68.12%. Remarkably, we have gathered functional evidence that supports the involvement of these viroid-encoded peptides in the life cycle of HSVd and ELVd, representative members of families *Pospiviroidae* and *Avsunviroidae*, respectively. The easiness of cloning viroid sequences using the aforementioned strategy facilitated the obtention of the constructs harbouring the point

mutations affecting (stop) or not (silent mutation) the putative encoded peptides, especially since the amplification of dimeric cDNA templates using a high-fidelity polymerase with self-complementary mutagenic primers (Sanjuán and Daròs, 2007) occasionally resulted in frequent undesired mutations. The bioassays with the viroid variants carrying a stop codon when compared to their controls indicated that these mutations diminished the viroid biological efficiency. This result, together with the coincidence of the subcellular localization of the conserved peptides present in HSVd and ELVd with the organelle in which viroid replication/accumulation occurs (nucleolus and chloroplast, respectively) and the physical interaction with the translational machinery evidenced by polysome isolation constitute strong indirect evidence pointing towards the existence and functional relevance of such peptides. If true, that would suppose a paradigm shift in which viroids, the first circular RNAs described in nature, would move from the non-coding categorization, as it has happened for certain eukaryotic circular RNAs (Diallo et al., 2019; Patop et al., 2019). However, it cannot be ruled out that some viroids might be coding and others do not, since our research has been focused on just two viroids and there are more than 30 species described to date (Chiumenti et al., 2021). Considering that classic nuclear localization signals are usually composed of one or two clusters of basic amino acids, an argument for the universality of the phenomenon could be the fact that relatively rich regions in basic amino acids (KR) as such identified in HSVd sequence (Gómez and Pallás, 2007b) are present in the other nuclear-replicating viroids (Figure D.1A). In the same direction, we have predicted by computational methods the existence of chloroplast-localization signals in viroid-derived peptides of the other three members of the family *Avsunviroidae* (Figure D.1B). It is worth noting that the accession of CChMVd (DQ450682) in Figure D.1B is the same that was fused to a fluorescent protein and fluorescence signal was observed in the chloroplasts (Baek et al., 2017) although the authors attributed their results to organelle RNA transport mechanisms driven by viroid RNA sequences.

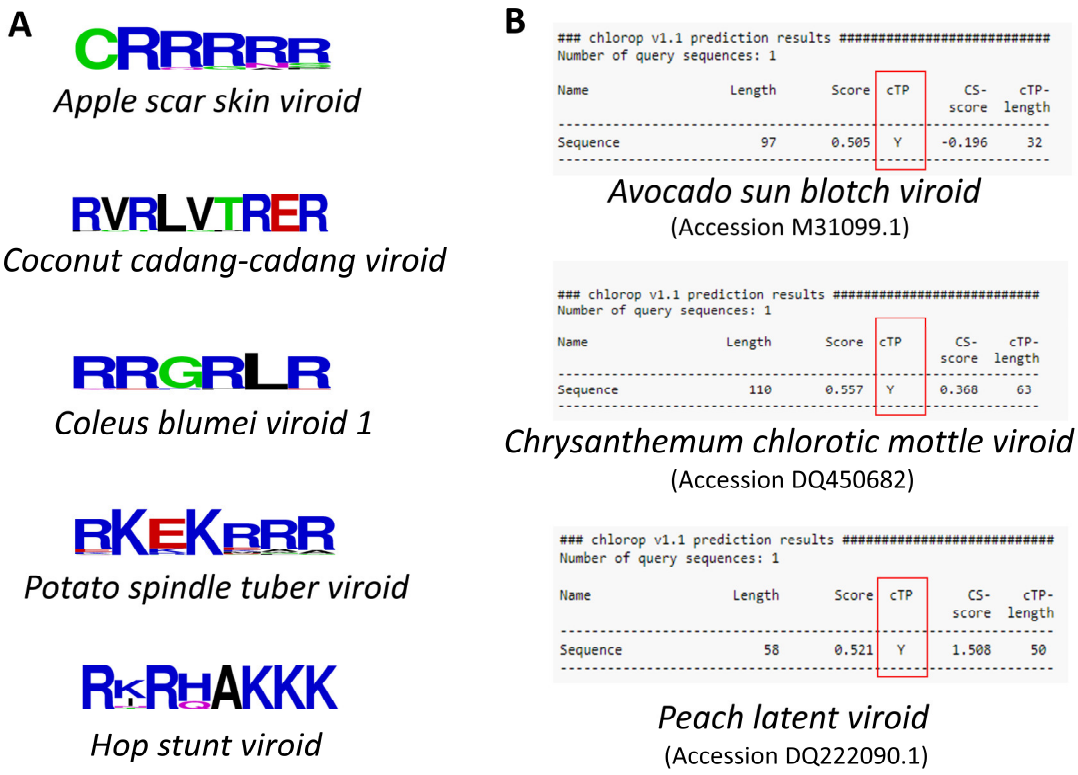


Figure D.1. A) K/R domains identified in the predicted peptides of the type species of each of the five genera that constitute the family *Pospiviroidae*. **B**) Computational prediction of chloroplast localization signals of peptides derived from viroids of the family *Avsunviroidae*. Length: is the length of the submitted sequence. Score is the output score from the second step network. The prediction cTP/no cTP is based solely on this score. cTP refers to whether or not this is predicted as a cTP-containing sequence; "Y" means that the sequence is predicted to contain a cTP; "-" means that is predicted not to contain a cTP. CS-score is the MEME scoring matrix score for the suggested cleavage site. The accession number of the viroid sequence in the nucleotide database of NCBI is indicated in brackets.

A subsequent speculation would be the functionality that putative viroid peptides may have. It is tempting to propose that could be related with viroid intracellular transport, a shady aspect of viroid biology (Pallás and Gómez, 2017). For this, the putative peptides should be able to bind viroid RNA or, at least, a host RNA-binding factor. Therefore, *in vivo* identification of ribonucleoprotein complexes containing viroid RNAs could be an interesting subject of further studies to solve this issue. Regarding ELVd-derived peptides, we have shown that they are proteolytically processed after chloroplast translocation but, as described for other signal peptides, most likely they are also degraded by the stromal transit peptide subfragment-degrading enzyme (Richter and Lamppa, 2002). Therefore, it seems unlikely that these peptides could have a functional role inside the chloroplast. It would be more plausible that ELVd-derived peptides interact with the TIC/TOC complexes to facilitate the transport of the viroid to the stroma through the inner and outer membranes, respectively. Since all the necessary RNAs in the chloroplast have been retained in the genome of this organelle (Barbrook et al., 2006), it is difficult to argue from an evolutionary point of view that a specific RNA transport mechanism independent of the

TIC/TOC complexes may exist (Gómez and Pallás, 2010c). In this line, foreign RNAs that enter into the plastids (chloroplast replicating viroids) may co-opt the only entry mechanism that has been described: the translocon machinery. The exit from the plastid may be triggered by accidental diffusion of viroid molecules through the translocon or mediated by a ribonucleoprotein complex that is relocated by different organelle export mechanisms (Krupinska et al., 2020). On the other hand, nucleocytoplasmic transport has been much studied (Tamura and Hara-Nishimura, 2014). Peptides targeted to the nucleus/nucleolus are not cleaved in the transport process, so additional functions besides facilitating the import could be performed by interacting with nuclear/nucleolar host factors. The nucleolus in particular has diverse functional roles (Kalinina et al., 2018) and it is targeted by many types of viruses (Taliansky et al., 2010). For example, some viruses hijack the nucleolar protein fibrillarin for systemic or cell-to-cell RNA trafficking (Kim et al., 2007; Sang et al., 2007; Zheng et al., 2015; Li et al., 2018) and even the helper-virus independent long-distance trafficking of a satellite RNA requires this protein (Chang et al., 2016).

The presence of viroids in ribosomal fractions was previously shown for CEVd in tomato generating alterations in the global polysome profiles and ribosome biogenesis (Cottilli et al., 2019). These results together with those obtained in chapter two showing HSVd and ELVd interaction with the translational machinery indicate that, from a mechanistic point of view, it would make sense that viroids enter the cell through the plasmodesmata and, in the cytoplasm, interact with the ribosomes to be translated. Then, the resulting peptides could facilitate the compartmentalization of these RNAs, which is required for protecting the replication intermediates from the RNA silencing machinery.

Another interesting observation is that both ELVd and HSVd contain the consensus m6A methylation motif, and we detected m6A modification in HSVd, suggesting that it might be driving translation initiation as it has been shown for human circRNAs (Yang et al., 2017). The study of RNA modifications in viroid sequences is still a very unexplored aspect, as only the absence of m5C had been confirmed in two representative viroids (Di Serio et al., 2019).

Despite all this body of evidence indirectly supporting viroid translation, we were unable to detect the predicted viroid peptides in protein extracts from infected plants by mass spectrometry. Perhaps they are very short-lived peptides that undergo a fast turnover (as most likely happens in chloroplastic ones) or exist only at very specific steps in the viroid life cycle, both situations resulting in such a low accumulation level which could make the peptides undetectable by conventional methods. Therefore, more sensitive approaches for peptide detection may be either required to shed light on that aspect and to expand the translatome to circular RNAs in plants.

Finally, we have leveraged on three -omic technologies, sRNA-seq, RNA-seq and whole-genome bisulfite sequencing (WGBS), for a detailed characterization of HSVd infection and the changes that this pathogenic viroid triggers in cucumber that eventually lead to the stunting symptoms. For the analysis, we took three biological replicates at three time points, 10-, 17- and 24-days post-inoculation (dpi) that comprise from the initial infection stage in which HSVd reaches the systemic tissue and starts replicating (HSVd was detected in 2 out of 3 biological replicates at 10 dpi) to the symptomatic stage in which stunting symptoms are obvious (24 dpi). From the HSVd levels of each replicate, it can be inferred that viroid accumulation reaches a maximum and since then starts to decline. This is revealed by the observation that plants with the highest viroid levels at 10 dpi had the lowest ones at 24 dpi, while for the sample in which HSVd was not detected at 10 dpi, as the infection began later, the highest accumulation was observed at 24 dpi. This is in accordance with previous reports that suggest a diminution in viroid accumulation in late infection (Adkar-Purushothama and Perreault, 2020). It is unclear whether this apparent threshold to viroid accumulation might be mostly due to the activation of plant defences or conversely due to an exhaustion of the host machinery caused by the infection. The progression of viroid infection studied here had not been considered in next-generation sequencing experiments before, as most of the studies have been focused only on the analysis of symptomatic plants (Owens et al., 2012; Xia et al., 2017; Zheng et al., 2017; Thibaut and Claude, 2018; Góra-Sochacka et al., 2019; Štajner et al., 2019; Takino et al., 2019; Wang et al., 2019b), and therefore the alterations in early infection stages were mostly unreported.

The analysis of HSVd-sRNAs during the three time points showed a similar pattern, which is not surprising since similar profiles were even found across different viroid strains (Tsushima et al., 2015) and host cultivars (Wang et al., 2011; Zhang et al., 2020). Interestingly, of the 7428 unique HSVd-sRNAs, the three most accumulated represented 23.34% of the total reads, while the ten most abundant accounted for more than a third (35.31%). It will be interesting to determine if these extreme accumulation differences are due to processing hot-spots in the viroid molecules, as it has been previously attributed to these hyper-accumulating regions being more susceptible to RNA silencing (Wang et al., 2011; Tsushima et al., 2015) or differences in the stability and/or movement of the vd-sRNA due to the interaction with host factors. Indeed, a region of HSVd seems to be mostly inaccessible to the RNA silencing machinery, because of the lack of vd-sRNAs originating from that region despite our high sequencing depth. Thus, it seems plausible that there are differences in the DICER accessibility (maybe due to secondary structures and/or the interaction with other proteins) that may have an impact on the generation of vd-sRNAs. Nonetheless, a surprising fact that we observed is that the proportion of HSVd small RNAs of 24-nt in length decreased in favour of those of 21-nt over the infection. In fact, it was previously reported that the

predominant size of HSVd sRNAs in cucumber leaves was 21-nt (Martinez et al., 2010; Zhang et al., 2020) but that was reported at late infection stages. Therefore, according to our results, the accumulation of viroid small RNAs seems a dynamic phenomenon. In fact, this shift in the proportion of HSVd sRNAs, together with their presence in the phloem (Martinez et al., 2010) points towards the selective movement of HSVd-sRNAs that would result in accumulation biases. Another astounding evidence of accumulation biases is that, regardless of the infection time-point, two sizes of vd-sRNAs (20- and 23-nt) were statistically more abundant from the (−) strand. Additionally, the number of unique vd-sRNAs was not significantly different depending on the strand of origin which is a strong support for the origin of vd-sRNAs from replication intermediates (Gómez et al., 2009), especially since the (−) strand does not exist as a mature circular form (Branch et al., 1981).

Regarding host alterations, this comprehensive analysis has provided a global overview of the nature and intensity of the changes that occur during HSVd infection. At the initial stage (10 dpi), there is a preponderance of differential exon usage, which is almost anecdotic at 17 dpi and slightly increases at 24 dpi. This may be interpreted as a swift mechanism to expand the functional diversity of proteins which could be related to defense mechanisms (Mastrangelo et al., 2012), being that coherent with the fact that most of the differential exon usage (DEU) events at 10 dpi were the increased retention of exons. Transcriptional alterations at this initial stage were mostly the up-regulation of genes and, as the infection progresses, transcriptional downregulation was accentuated. Despite some of these transcripts could be directly silenced by small RNAs of the viroid (we found three by degradome analysis and validated two of them by a reporter assay), this must represent only a very small fraction of the genes that decrease their expression as a consequence of the infection. Still, the silencing of these particular genes might be beneficial for viroid spread/accumulation in coincidence with what has been observed in other viroids (discussed in chapter 3). Therefore, although it is unquestionable that, similarly to viruses, RNA silencing is linked with viroid symptoms (Ramesh et al., 2021), it fails to explain the whole picture and might not be the trigger of the initial molecular lesion in nuclear-replicating viroids (Flores et al., 2020).

In our study, the functional activity of 24-nt vd-sRNA in the RdDM pathway was very limited, as we only found two partially homologous (2 mismatches) to the cucumber genome, and just one affected the methylation status causing hypermethylation (CHG context) in an intron of CsaV3_4G027560. However, no transcriptional differences in the expression of this gene were observed. This confirms that 24-nt vd-sRNAs are incorporated into the RdDM pathway (Wassenegger and Dalakouras, 2021), but lack physiological relevance for HSVd-cucumber pathogenesis, and probably for other host-viroid combinations also. Supporting this hypothesis, the

homologies found between group I introns (found in nuclear rRNA genes) and the CCR of nuclear-replicating viroids are well below 20nt (Dinter-Gottlieb, 1986) and thus would discard a functional involvement.

Therefore, the methylation changes that we have detected through the infection should be related either to the activation of plant defense mechanisms in response to the pathogen and/or the subversion of host proteins by the viroid (as is the case of HDA6). It is curious that, although the differences were very subtle for the whole genome, the dynamics were coincident with the ribosomal gene: early hypomethylation that turned into hypermethylation at the late infection. The rRNA genes are firstly hypomethylated as a consequence of HSVd infection, but, although extended to the whole unit, this primarily occurs in the promoter region, which is compatible with the initial involvement of HDA6 (Castellano et al., 2016b) that later is probably counteracted by another endogenous mechanism to recover the homeostasis. Given that this hypomethylation is detected at 10 dpi, when no HSVd-transcripts (presumably replication intermediates) were recovered for a biological replicate, it seems that this alteration precedes, and could be a requirement for, viroid replication, as it was previously proposed (Castellano et al., 2016b).

Viroids, as naked RNAs, might initially dodge some plant defense mechanisms, and only upon replication, they trigger RNA silencing and interact with host proteins producing imbalances that are later recognized as symptoms. Maybe because of this, viroid symptoms such as stunting are not caused as a simple perturbation but complex alterations that require an integrative analysis to understand the molecular basis of the process.

CONCLUSIONS

- The use of type IIS restrictions enzymes facilitates the dimerization of viroid sequences and enables the direct and oriented cloning into a binary vector, constituting a reliable and efficient method to obtain an infectious clone of any viroid.

- HSVd and ELVd contain conserved ORFs within their corresponding variants able to encode functional peptides carrying specific subcellular localization signals that direct fluorescent-fused proteins to the corresponding organelle where replication/accumulation takes place.

- Mutations that truncate the nucleolar domain of HSVd are detrimental for the viroid while truncating any of the two ELVd ORF that contains a cTP also diminishes (but to a lesser extent) viroid biological efficiency, maybe because of functional redundancy.

- Circular forms of both, HSVd and ELVd RNAs were found in polysome fractions, revealing their physical interaction with the translational machinery of the plant cell.

- HSVd and ELVd contain the consensus m6A methylation motif RRACH, and we detected m6A modification in HSVd, suggesting that it might be driving translation initiation as it has been shown for human circRNAs.

- HSVd-sRNAs in cucumber are predominantly 24-nt in length, although the proportion of 24-nt diminishes in favour of those of 21-nt with the progression of the infection. Approximately 40% of the small RNAs of HSVd correspond to four hyper-accumulating peaks and most HSVd-sRNAs are derived from the plus strand except those of 20- and 23-nt in length. Considering unique sequences, no differences between the plus and the minus strand were observed, and one region of the viroid seems inaccessible to the silencing machinery. Taken together, these results revealed important accumulation biases of vd-sRNAs which might be due to the processing of hot-spots or to differences in the stability or movement.

- The functionality of HSVd-sRNAs is restricted to the silencing of certain transcripts: we found evidence for RTM2 and AQP. Silencing these transcripts might benefit the viroid in agreement with the targets described for other nuclear-replicating viroids.

- HSVd triggers systemic differential exon usage mainly at early infection and transcriptional repression over time in cucumber genes. Among the most prevalent functional annotations, transport and membrane were the most enriched. The alterations regarding miRNAs and other small RNAs are quite limited and mainly occur at the late infection.

- HSVd reaches a systemic infection on cucumber at 10 dpi causing the hypomethylation of the ribosomal genes, which is more noticeable in the promoter region. This hypomethylation turns into hypermethylation as the infection progresses which coincides with the general trend of the whole genome. Only the methylation differences located in promoter or CDS correlate with transcriptional alterations.

REFERENCES

AbouHaidar MG, Venkataraman S, Golshani A, Liu B, Ahmad T (2014) Novel coding, translation, and gene expression of a replicating covalently closed circular RNA of 220 nt. Proc Natl Acad Sci U S A **111**: 14542–14547

Adkar-Purushothama CR, Brosseau C, Giguère T, Sano T, Moffett P, Perreault J-P (2015) Small RNA Derived from the Virulence Modulating Region of the Potato spindle tuber viroid Silences callose synthase Genes of Tomato Plants. Plant Cell **27**: 2178–94

Adkar-Purushothama CR, Iyer PS, Perreault J-P (2017) Potato spindle tuber viroid infection triggers degradation of chloride channel protein CLC-b-like and Ribosomal protein S3a-like mRNAs in tomato plants. Sci Rep **7**: 8341

Adkar-Purushothama CR, Perreault J-P (2018) Alterations of the viroid regions that interact with the host defense genes attenuate viroid infection in host plant. RNA Biol **15**: 955–966

Adkar-Purushothama CR, Perreault JP (2020) Current overview on viroid–host interactions. Wiley Interdiscip Rev RNA **11**: e1570

Adkar-Purushothama CR, Sano T, Perreault JP (2018) Viroid-derived small RNA induces early flowering in tomato plants by RNA silencing. Mol Plant Pathol **19**: 2446–2458

Afzal Z, Howton T, Sun Y, Mukhtar M (2016) The Roles of Aquaporins in Plant Stress Responses. J Dev Biol **4**: 9

Amari K, Gomez G, Myrta A, Di Terlizzi B, Pallás V (2001) The molecular characterization of 16 new sequence variants of *Hop stunt viroid* reveals the existence of invariable regions and a conserved hammerhead-like structure on the viroid molecule. J Gen Virol **82**: 953–962

Anders S, Pyl PT, Huber W (2015) HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics **31**: 166–169

Anders S, Reyes A, Huber W (2012) Detecting differential usage of exons from RNA-seq data. Genome Res **22**: 2008–2017

Annacondia ML, Magerøy MH, Martinez G (2018) Stress response regulation by epigenetic mechanisms: changing of the guards. Physiol Plant **162**: 239–250

Apostolova E, Hadjieva N, Ivanova DP, Yahubyan G, Baev V, Gozmanova M (2021) MicroRNA expression dynamics reshape the cultivar-specific response of pepper (Capsicum annuum L.) to potato spindle tuber viroid (PSTVd) infection. Sci Hortic (Amsterdam) **278**: 109845

Arribas-Hernández L, Brodersen P (2020) Occurrence and functions of m6A and other covalent modifications in plant mRNA. Plant Physiol **182**: 79–96

Avina-Padilla K, Martinez de la Vega O, Rivera-Bustamante R, Martinez-Soriano JP, Owens RA, Hammond RW, Vielle-Calzada J-P (2015) In silico prediction and validation of potential gene targets for pospiviroid-derived small RNAs during tomato infection. Gene **564**: 197–205

Axtell MJ (2013) Classification and Comparison of Small RNAs from Plants. Annu Rev Plant Biol **64**: 137–159

Baek E, Park M, Yoon J-Y, Palukaitis P (2017) Chrysanthemum Chlorotic Mottle Viroid-Mediated Trafficking of Foreign mRNA into Chloroplasts. Res Plant Dis **23**: 288–293

Bao S, Owens RA, Sun Q, Song H, Liu Y, Eamens AL, Feng H, Tian H, Wang M-B, Zhang R (2019) Silencing of transcription factor encoding gene StTCP23 by small RNAs derived from the virulence modulating region of potato spindle tuber viroid is associated with symptom development in potato. PLOS Pathog **15**: e1008110

Barbrook AC, Howe CJ, Purton S (2006) Why are plastid genomes retained in non-photosynthetic organisms? Trends Plant Sci **11**: 101–108

Barchi L, Pietrella M, Venturini L, Minio A, Toppino L, Acquadro A, Andolfo G, Aprea G, Avanzato C, Bassolino L, and others (2019) A chromosome-anchored eggplant genome sequence reveals key events in Solanaceae evolution. Sci Rep **9**: 11769

Bardou F, Ariel F, Simpson CG, Romero-Barrios N, Laporte P, Balzergue S, Brown JWS, Crespi M (2014) Long Noncoding RNA Modulates Alternative Splicing Regulators in Arabidopsis. Dev Cell **30**: 166–176

Baulcombe D (2004) RNA silencing in plants. Nature **431**: 356–363

Baulcombe DC, Dean C (2014) Epigenetic regulation in plant responses to the environment. Cold Spring Harb Perspect Biol **6**: a019471–a019471

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B **57**: 289–300

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114–2120

Bonfiglioli RG, Webb DR, Symons RH (1996) Tissue and intra-cellular distribution of coconut cadang cadang viroid and citrus exocortis viroid determined by in situ hybridization and confocal laser scanning and transmission electron microscopy. Plant J **9**: 457–465

Bouché N, Lauressergues D, Gasciolli V, Vaucheret H (2006) An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. EMBO J **25**: 3347–3356

Branch A, Robertson H (1984) A replication cycle for viroids and other small infectious RNA's. Science (80-) **223**: 450–455

Branch AD, Robertson HD, Dickson E (1981) Longer-than-unit-length viroid minus strands are present in RNA from infected plants. Proc Natl Acad Sci **78**: 6381–6385

Brodersen P, Voinnet O (2006) The diversity of RNA silencing pathways in plants. Trends Genet **22**: 268–280

Bruce BD (2000) Chloroplast transit peptides: structure, function and evolution. Trends Cell Biol **10**: 440–447

Carbonell A, De la Peña M, Flores R, Gago S (2006) Effects of the trinucleotide preceding the self-cleavage site on eggplant latent viroid hammerheads: differences in co- and post-transcriptional self-cleavage may explain the lack of trinucleotide AUC in most natural hammerheads. Nucleic Acids Res **34**: 5613–22

Carbonell A, Martínez de Alba ÁE, Flores R, Gago S (2008) Double-stranded RNA interferes in a sequence-specific manner with the infection of representative members of the two viroid families. Virology **371**: 44–53

Carbonell A, Takeda A, Fahlgren N, Johnson SC, Cuperus JT, Carrington JC (2014) New generation of artificial MicroRNA and synthetic trans-acting small interfering RNA vectors for efficient gene silencing in Arabidopsis. Plant Physiol **165**: 15–29

Carmo-Fonseca M, Mendes-Soares L, Campos I (2000) To be or not to be in the nucleolus. Nat Cell Biol **2**: 107–112

Castellano M, Martinez G, Marques MC, Moreno-Romero J, Köhler C, Pallas V, Gomez G (2016a) Changes in the DNA methylation pattern of the host male gametophyte of viroid-infected cucumber plants. J Exp Bot **67**: 5857–5868

Castellano M, Martinez G, Pallás V, Gómez G (2015) Alterations in host DNA methylation in response to constitutive expression of Hop stunt viroid RNA in Nicotiana benthamiana plants. Plant Pathol **64**: 1247–1257

Castellano M, Pallas V, Gomez G (2016b) A pathogenic long noncoding RNA redesigns the epigenetic landscape of the infected cells by subverting host Histone Deacetylase 6 activity. New Phytol **211**: 1311–1322

Catoni M, Tsang JMF, Greco AP, Zabet NR (2018) DMRcaller: A versatile R/Bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. Nucleic Acids Res **46**: e114–e114

Cech TR, Zaug AJ, Grabowski PJ (1981) In vitro splicing of the ribosomal RNA precursor of tetrahymena: Involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell **27**: 487–496

Cervera A, De La Peña M (2020) Small circRNAs with self-cleaving ribozymes are highly expressed in diverse metazoan transcriptomes. Nucleic Acids Res **48**: 5054–5064

Chang C-H, Hsu F-C, Lee S-C, Lo Y-S, Wang J-D, Shaw J, Taliansky M, Chang B-Y, Hsu Y-H, Lin N-S (2016) The Nucleolar Fibrillarin Protein Is Required for Helper Virus-Independent Long-Distance Trafficking of a Subviral Satellite RNA in Plants. Plant Cell **28**: 2586–2602

Charrad M, Ghazzali N, Boiteau V, Niknafs A (2014) Nbclust: An R package for determining the relevant number of clusters in a data set. J Stat Softw **61**: 1–36

Chen C, Li J, Feng J, Liu B, Feng L, Yu X, Li G, Zhai J, Meyers BC, Xia R (2021) sRNAanno—a database repository of uniformly annotated small RNAs in plants. Hortic Res 2021 81 **8**: 1–8

Chen L, Zhang P, Fan Y, Lu Q, Li Q, Yan J, Muehlbauer GJ, Schnable PS, Dai M, Li L (2018a) Circular RNAs mediated by transposons are associated with transcriptomic and phenotypic variation in maize. New Phytol **217**: 1292–1306

Chen S, Zhou Y, Chen Y, Gu J (2018b) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics **34**: i884–i890

Chiumenti M, Navarro B, Candresse T, Flores R, Di Serio F (2021) Reassessing species demarcation criteria in viroid taxonomy by pairwise identity matrices. Virus Evol **7**: 1–14

Codoñer FM, Daròs JA, Solé R V., Elena SF (2006) The fittest versus the flattest: Experimental confirmation of the quasispecies effect with subviral pathogens. PLoS Pathog **2**: 1187–1193

Conn VM, Hugouvieux V, Nayak A, Conos SA, Capovilla G, Cildir G, Jourdain A, Tergaonkar V, Schmid M, Zubieta C, and others (2017) A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation. Nat Plants **3**: 17053

Cordero T, Ortolá B, Daròs JA (2018) Mutational analysis of Eggplant Latent Viroid RNA circularization by the eggplant tRNA ligase in Escherichia coli. Front Microbiol **9**: 635

Corrêa RL, Sanz-Carbonell A, Kogej Z, Müller SY, Ambrós S, López-Gomollón S, Gómez G, Baulcombe DC, Elena SF (2020) Viral Fitness Determines the Magnitude of Transcriptomic and Epigenomic Reprograming of Defense Responses in Plants. Mol Biol Evol **37**: 1866–1881

Cottilli P, Belda-Palazón B, Adkar-Purushothama CR, Perreault JP, Schleiff E, Rodrigo I, Ferrando A, Lisón P (2019) Citrus exocortis viroid causes ribosomal stress in tomato plants. Nucleic Acids Res **47**: 8649–8661

Couso J-P, Patraquim P (2017) Classification and function of small open reading frames. Nat Rev Mol Cell Biol 2017 189 **18**: 575–589

Cress DE, Kiefer MC, Owens RA (1983) Construction of infectious potato spindle tuber viroid cDNA clones. Nucleic Acids Res **11**: 6821–35

Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: A Sequence Logo Generator. Genome Res **14**: 1188–1190

Csorba T, Kontra L, Burgyán J (2015) Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. Virology **479**–**480**: 85–103

Dadami E, Boutla A, Vrettos N, Tzortzakaki S, Karakasilioti I, Kalantidis K (2013) DICER-LIKE 4 But Not DICER-LIKE 2 may have a positive effect on potato spindle tuber viroid accumulation in nicotiana benthamiana. Mol Plant **6**: 232–234

Dai X, Zhuang Z, Zhao PX (2018) psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res **46**: W49–W54

Dalakouras A, Dadami E, Bassler A, Zwiebel M, Krczal G, Wassenegger M (2015) Replicating potato spindle tuber viroid mediates de novo methylation of an intronic viroid sequence but no cleavage of the corresponding pre-mRNA. RNA Biol **12**: 268–275

Dalakouras A, Dadami E, Wassenegger M (2013) Viroid-induced DNA methylation in plants. Biomol Concepts **4**: 557–565

Dalakouras A, Moser M, Zwiebel M, Krczal G, Hell R, Wassenegger M (2009) A hairpin RNA construct residing in an intron efficiently triggered RNA-directed DNA methylation in tobacco. Plant J **60**: 840–851

Dang T, Lavagi-Craddock I, Bodaghi S, Vidalakis G (2021) Next-Generation Sequencing Identification and Characterization of MicroRNAs in Dwarfed Citrus Trees Infected With Citrus Dwarfing Viroid in High-Density Plantings. Front Microbiol **12**: 980

Darbani B, Noeparvar S, Borg S (2016) Identification of circular RNAs from the parental genes involved in multiple aspects of cellular metabolism in barley. Front Plant Sci **7**: 776

Daròs JA (2016) Eggplant latent viroid: a friendly experimental system in the family Avsunviroidae. Mol Plant Pathol **17**: 1170–1177

Daròs JA, Aragonés V, Cordero T (2018) A viroid-derived system to produce large amounts of recombinant RNA in Escherichia coli. Sci Rep **8**: 1–9

Daròs JA, Elena SF, Flores R (2006) Viroids: An Ariadne's thread into the RNA labyrinth. EMBO Rep **7**: 593–598

Daròs JA, Marcos JF, Hernández C, Flores R (1994) Replication of avocado sunblotch viroid: evidence for a symmetric pathway with two rolling circles and hammerhead ribozyme processing. Proc Natl Acad Sci U S A **91**: 12813–7

Davies JW, Kaesberg P, Diener TO (1974) Potato spindle tuber viroid. XII. An investigation of viroid RNA as a messenger for protein synthesis. Virology **61**: 281–286

Deleris A, Halter T, Navarro L (2016) DNA Methylation and Demethylation in Plant Immunity. Annu Rev Phytopathol **54**: 579–603

Delgado S, Navarro B, Serra P, Gentit P, Cambra MÁ, Chiumenti M, De Stradis A, Di Serio F, Flores R (2019) How sequence variants of a plastid-replicating viroid with one single nucleotide change initiate disease in its natural host. RNA Biol **16**: 906–917

Denti MA, Boutla A, Tsagris M, Tabler M (2004) Short interfering RNAs specific for potato spindle tuber viroid are found in the cytoplasm but not in the nucleus. Plant J **37**: 762–9

Diallo LH, Tatin F, David F, Godet A-C, Zamora A, Prats A-C, Garmy-Susini B, Lacazette E (2019) How are circRNAs translated by non-canonical initiation mechanisms? Biochimie **164**: 45–52

Diener TO (1989) Circular RNAs: Relics of precellular evolution? Proc Natl Acad Sci U S A **86**: 9370–9374

Diener TO (1971) Potato spindle tuber "virus": IV. A replicating, low molecular weight RNA. Virology **45**: 411–428

Diener TO (2003) Discovering viroids: a personal perspective. Nat Rev Microbiol **1**: 75–80

Diermann N, Matoušek J, Junge M, Riesner D, Steger G (2010) Characterization of plant miRNAs and small RNAs derived from potato spindle tuber viroid (PSTVd) in infected tomato. **391**: 1379–1390

Ding B (2009) The Biology of Viroid-Host Interactions. Annu Rev Phytopathol **47**: 105–131

Ding B, Itaya A (2007) Viroid: A useful model for studying the basic principles of infection and RNA biology. Mol Plant-Microbe Interact **20**: 7–20

Ding B, Wang Y (2009) Viroids: Uniquely Simple and Tractable Models to Elucidate Regulation of Cell-to-Cell Trafficking of RNA. DNA Cell Biol **28**: 51–56

Ding SW, Voinnet O (2007) Antiviral Immunity Directed by Small RNAs. Cell **130**: 413–426

Dingley AJ, Grzesiek S (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 2J(NN) couplings. J Am Chem Soc **120**: 8293–8297

Dingley AJ, Steger G, Esters B, Riesner D, Grzesiek S (2003) Structural characterization of the 69 nucleotide potato spindle tuber viroid left-terminal domain by NMR and thermodynamic analysis. J Mol Biol **334**: 751–767

Dinter-Gottlieb G (1986) Viroids and virusoids are related to group I introns. Proc Natl Acad Sci U S A **83**: 6250–6244

Di Serio F, Flores R, Verhoeven JTJ, Li S-F, Pallás V, Randles JW, Sano T, Vidalakis G, Owens RA (2014) Current status of viroid taxonomy. Arch Virol **159**: 3467–3478

Di Serio F, Gisel A, Navarro B, Delgado S, de Alba ÁEM, Donvito G, Flores R (2009) Deep sequencing of the small RNAs derived from two symptomatic variants of a chloroplastic viroid: Implications for their genesis and for pathogenesis. PLoS One **4**: e7539

Di Serio F, Li S-F, Pallás V, Owens RA, Randles JW, Sano T, Verhoeven JTJ, Vidalakis G, Flores R (2017) Chapter 13 - Viroid Taxonomy. *In* A Hadidi, R Flores, JW Randles, PBT-V and S Palukaitis, reds,Academic Press, Boston, bll 135–146

Di Serio F, Li SF, Matoušek J, Owens RA, Pallás V, Randles JW, Sano T, Verhoeven JTJ, Vidalakis G, Flores R (2018) ICTV virus taxonomy profile: Avsunviroidae. J Gen Virol **99**: 611–612

Di Serio F, Martinez de Alba A-E, Navarro B, Gisel A, Flores R (2010) RNA-Dependent RNA Polymerase 6 Delays Accumulation and Precludes Meristem Invasion of a Viroid That Replicates in the Nucleus. J Virol **84**: 2477–2489

Di Serio F, Owens RA, Li SF, Matoušek J, Pallás V, Randles JW, Sano T, Verhoeven JTJ, Vidalakis G, Flores R (2021) ICTV virus taxonomy profile: Pospiviroidae. J Gen Virol **102**: 1543

Di Serio F, Torchetti EM, Daròs JA, Navarro B (2019) Reassessment of viroid RNA cytosine methylation status at the single nucleotide level. Viruses **11**: 357

de la Peña M, García-Robles I, Cervera A (2017) The Hammerhead Ribozyme: A Long History for a Short RNA. Molecules **22**: 78

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29**: 15–21

Dunoyer P, Himber C, Voinnet O (2005) DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. Nat Genet **37**: 1356–1360

Dynowski M, Schaaf G, Loque D, Moran O, Ludewig U (2008) Plant plasma membrane water channels conduct the signalling molecule H2O2. Biochem J **414**: 53–61

Eamens AL, Smith NA, Dennis ES, Wassenegger M, Wang MB (2014) In Nicotiana species, an artificial microRNA corresponding to the virulence modulating region of Potato spindle tuber viroid directs RNA silencing of a soluble inorganic pyrophosphatase gene and the development of abnormal phenotypes. Virology **450**–**451**: 266–277

Earley KW, Pontvianne F, Wierzbicki AT, Blevins T, Tucker S, Costa-Nunes P, Pontes O, Pikaard CS (2010) Mechanisms of HDA6-mediated rRNA gene silencing: Suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. Genes Dev **24**: 1119–1132

Elena SF, Gómez G, Daròs JA (2009) Evolutionary constraints to viroid evolution. Viruses **1**: 241–254

Engler C, Gruetzner R, Kandzia R, Marillonnet S (2009) Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. PLoS One **4**: e5553

Fadda Z, Daròs JA, Fagoaga C, Flores R, Duran-Vila N (2003) Eggplant Latent Viroid, the Candidate Type Species for a New Genus within the Family Avsunviroidae (Hammerhead Viroids). J Virol **77**: 6528–6532

Fagoaga C, Pina JA, Duran-Vila N (1994) Occurrence of small RNAs in severely diseased vegetable crops. Plant Dis **78**: 749–753

Fan J, Quan W, Li G-B, Hu X-H, Wang Q, Wang H, Li X-P, Luo X, Feng Q, Hu Z-J, and others (2020) circRNAs Are Involved in the Rice Magnaporthe oryzae Interaction. Plant Physiol **182**: 272–286

Fang Y, Zheng Y, Lu W, Li J, Duan Y, Zhang S, Wang Y (2021) Roles of miR319-regulated TCPs in plant development and response to abiotic stress. Crop J **9**: 17–28

Fares MA (2015) The origins of mutational robustness. Trends Genet **31**: 373–381

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. Nature **391**: 806–811

Flores R, Delgado S, Gas M-E, Carbonell A, Molina D, Gago S, De la Peña M (2004) Viroids: the minimal non-coding RNAs with autonomous replication. FEBS Lett **567**: 42–48

Flores R, Delgado S, Rodio ME, Ambrós S, Hernández C, Di Serio F (2006) Peach latent mosaic viroid: Not so latent. Mol Plant Pathol **7**: 209–221

Flores R, Gago-Zachert S, Serra P, Sanjuán R, Elena SF (2014) Viroids: Survivors from the RNA World? Annu Rev Microbiol **68**: 395–414

Flores R, Hernández C, Alba AEM de, Daròs J-A, Serio F Di (2005) Viroids and Viroid-Host Interactions. Annu Rev Phytopathol **43**: 117–139

Flores R, Minoia S, Carbonell A, Gisel A, Delgado S, López-Carrasco A, Navarro B, Di Serio F (2015) Viroids, the simplest RNA replicons: How they manipulate their hosts for being propagated and how their hosts react for containing the infection. Virus Res **209**: 136–145

Flores R, Navarro B, Delgado S, Serra P, Di Serio F (2020) Viroid pathogenesis: A critical appraisal of the role of RNA silencing in triggering the initial molecular lesion. FEMS Microbiol Rev **44**: 386–398

Flores R, Navarro JA, De La Peña M, Navarro B, Ambrós S, Vera A (1999) Viroids with hammerhead ribozymes: Some unique structural and functional aspects with respect to other members of the group. Biol Chem **380**: 849–854

Flores R, Owens RA, Taylor J (2016) Pathogenesis by subviral agents: Viroids and hepatitis delta virus. Curr Opin Virol **17**: 87–94

Gago-Zachert S (2016) Viroids, infectious long non-coding RNAs with autonomous replication. Virus Res **212**: 12–24

Gago S, Elena SF, Flores R, Sanjuán R (2009) Extremely High Mutation Rate of a Hammerhead Viroid. Science (80-) **323**: 1308–1308

Gardner RC, Chonoles KR, Owens RA (1986) Potato spindle tuber viroid infections mediated by the Ti plasmid of Agrobacterium tumefaciens. Plant Mol Biol **6**: 221–228

Gas M-E, Hernández C, Flores R, Daròs J-A (2007) Processing of Nuclear Viroids In Vivo: An Interplay between RNA Conformations. PLoS Pathog **3**: e182

Genoves A, Navarro JA, Pallás V, Genovés A, Navarro JA, Pallás V (2006) Functional analysis of the five melon necrotic spot virus genome-encoded proteins. J Gen Virol **87**: 2371–2380

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods **6**: 343–345

Giguère T, Adkar-Purushothama CR, Bolduc F, Perreault JP (2014a) Elucidation of the structures of all members of the Avsunviroidae family. Mol Plant Pathol **15**: 767–779

Giguère T, Adkar-Purushothama CR, Perreault J-P (2014b) Comprehensive secondary structure elucidation of four genera of the family Pospiviroidae. PLoS One **9**: e98655

Gómez-Díaz E, Jordà M, Peinado MA, Rivero A (2012) Epigenetics of Host–Pathogen Interactions: The Road Ahead and the Road Behind. PLoS Pathog **8**: e1003007

Gómez G, Martínez G, Pallás V (2008) Viroid-induced symptoms in Nicotiana benthamiana plants are dependent on RDR6 activity. Plant Physiol **148**: 414–423

Gómez G, Martínez G, Pallás V (2009) Interplay between viroid-induced pathogenesis and RNA silencing pathways. Trends Plant Sci **14**: 264–9

Gómez G, Pallás V (2001) Identification of an in vitro ribonucleoprotein complex between a viroid RNA and a phloem protein from cucumber plants. Mol Plant-Microbe Interact **14**: 910–913

Gómez G, Pallás V (2007a) Mature monomeric forms of Hop stunt viroid resist RNA silencing in transgenic plants. Plant J **51**: 1041–1049

Gómez G, Pallás V (2007b) A peptide derived from a single-modified viroid-RNA can be used as an "in vivo" nucleolar marker. J Virol Methods **144**: 169–171

Gómez G, Pallás V (2010a) Noncoding RNA mediated traffic of foreign mRNA into chloroplasts reveals a novel signaling mechanism in plants. PLoS One **5**: e12269

Gómez G, Pallás V (2004) A long-distance translocatable phloem protein from cucumber forms a ribonucleoprotein complex in vivo with Hop stunt viroid RNA. J Virol **78**: 10104–10

Gómez G, Pallás V (2006) Hop stunt viroid is processed and translocated in transgenic Nicotiana benthamiana plants. Mol Plant Pathol **7**: 511–517

Gómez G, Pallás V (2010b) Noncoding RNA mediated traffic of foreign mRNA into chloroplasts reveals a novel signaling mechanism in plants. PLoS One **5**: 1517–9

Gómez G, Pallás V (2010c) Can the import of mRNA into chloroplasts be mediated by a secondary structure of a small non-coding RNA? Plant Signal Behav **5**: 1517–9

Gómez G, Pallás V (2012a) Studies on subcellular compartmentalization of plant pathogenic noncoding RNAs give new insights into the intracellular RNA-traffic mechanisms. Plant Physiol **159**: 558–564

Gómez G, Pallás V (2013) Viroids: A light in the darkness of the lncRNA-directed regulatory networks in plants. New Phytol **198**: 10–15

Gómez G, Pallás V (2012b) A pathogenic non coding RNA that replicates and accumulates in chloroplasts traffics to this organelle through a nuclear-dependent step. Plant Signal Behav **7**: 881–883

Gómez G, Torres H, Pallás V (2004) Identification of translocatable RNA-binding phloem proteins from melon, potential components of the long-distance RNA transport system. Plant J **41**: 107–116

Góra-Sochacka A, Więsyk A, Fogtman A, Lirski M, Zagórski-Ostoja W (2019) Root Transcriptomic Analysis Reveals Global Changes Induced by Systemic Infection of Solanum lycopersicum with Mild and Severe Variants of Potato Spindle Tuber Viroid. Viruses **11**: 992

Gordon K, Fütterer J, Hohr T (1992) Efficient initiation of translation at non-AUG triplets in plant cells. Plant J **2**: 809–813

Gozmanova M, Denti MA, Minkov IN, Tsagris M, Tabler M (2003) Characterization of the RNA motif responsible for the specific interaction of potato spindle tuber viroid RNA (PSTVd) and the tomato protein Virp1. Nucleic Acids Res **31**: 5534–5543

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, and others (2011) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol **29**: 644

Gramazio P, Yan H, Hasing T, Vilanova S, Prohens J, Bombarely A (2019) Whole-Genome

Resequencing of Seven Eggplant (Solanum melongena) and One Wild Relative (S. incanum) Accessions Provides New Insights and Breeding Tools for Eggplant Enhancement. Front Plant Sci **10**: 1–17

Gross HJ, Domdey H, Lossow C, Jank P, Raba M, Alberty H, Sänger HL (1978) Nucleotide sequence and secondary structure of potato spindle tuber viroid. Nature **273**: 203–208

Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL (2008) The Vienna RNA Websuite. Nucleic Acids Res **36**: W70–W74

Guo JU, Agarwal V, Guo H, Bartel DP (2014) Expanded identification and characterization of mammalian circular RNAs. Genome Biol **15**: 409

Hall TC, Wepprich RK, Davies JW, Weathers LG, Semancik JS (1974) Functional distinctions between the ribonucleic acids from citrus exocortis viroid and plant viruses: Cell-free translation and aminoacylation reactions. Virology **61**: 486–492

Hamera S, Song X, Su L, Chen X, Fang R (2012) Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. Plant J **69**: 104–115

Hammann C, Steger G (2012) Viroid-specific small RNA in plant disease. RNA Biol **9**: 809–819

Hammond RW (2017) Economic Significance of Viroids in Vegetable and Field Crops. In: Viroids and Satellites. Chapter 1. Edited by A. Hadidi, R. Flores, J. W. Randles and P. Palukaitis. Academic Press, Elsevier. 754 pp. pp. 5-13 ISBN: 978-0-12-801498-1.

Han Y, Li X, Yan Y, Duan MH, Xu JH (2020) Identification, characterization, and functional prediction of circular RNAs in maize. Mol Genet Genomics **295**: 491–503

Hansen ER, Petracek ME, Dickey LF, Thompson WF (2001) The 5' End of the Pea Ferredoxin-1 mRNA Mediates Rapid and Reversible Light-Directed Changes in Translation in Tobacco. Plant Physiol **125**: 770–778

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J (2013) Natural RNA circles function as efficient microRNA sponges. Nature **495**: 384–388

Harders J, Lukács N, Robert-Nicoud M, Jovin TM, Riesner D (1989) Imaging of viroids in nuclei from tomato leaf tissue by in situ hybridization and confocal laser scanning microscopy. EMBO J **8**: 3941–9

Healey A, Furtado A, Cooper T, Henry RJ (2014) Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant Methods **10**: 1–8

Henco K, Riesner D, Sanger HL (1977) Conformation of viroids. Nucleic Acids Res **4**: 177–94

Henco K, Sänger HL, Riesner D (1979) Fine structure melting of viroids as studied by kinetic methods. Nucleic Acids Res **6**: 3041–59

Henras AK, Soudet J, Gérus M, Lebaron S, Caizergues-Ferrer M, Mougin A, Henry Y (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci **65**: 2334–2359

Herranz MC, Niehl A, Rosales M, Fiore N, Zamorano A, Granell A, Pallas V (2013) A remarkable synergistic effect at the transcriptomic level in peach fruits doubly infected by prunus necrotic ringspot virus and peach latent mosaic viroid. Virol J **10**: 164

Herranz MC, Sanchez-Navarro JA, Aparicio F, Pallás V (2005) Simultaneous detection of six

stone fruit viruses by non-isotopic molecular hybridization using a unique riboprobe or 'polyprobe'. J Virol Methods **124**: 49–55

Hirakawa H, Shirasawa K, Miyatake K, Nunome T, Negoro S, Ohyama A, Yamaguchi H, Sato S, Isobe S, Tabata S, and others (2014) Draft genome sequence of eggplant (solanum melongena L.): The representative solanum species indigenous to the old world. DNA Res **21**: 649–660

Hoffer P, Ivashuta S, Pontes O, Vitins A, Pikaard C, Mroczka A, Wagner N, Voelker T (2011) Posttranscriptional gene silencing in nuclei. Proc Natl Acad Sci U S A **108**: 409–414

Hou Y, Aydin E, De Bastiani M, Xiao C, Isikgor FH, Xue DJ, Chen B, Chen H, Bahrami B, Chowdhury AH, and others (2020) Efficient tandem solar cells with solution-processed perovskite on textured crystalline silicon. Science (80-) **367**: 1135–1140

Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, Kasschau KD, Carrington JC (2007) Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. Plant Cell **19**: 926–942

Hsu MT, Coca-Prados M (1979) Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells [24]. Nature **280**: 339–340

Hua XJ, Van De Cotte B, Van Montagu M, Verbruggen N (2001) The 5′ untranslated region of the At-P5R gene is involved in both transcriptional and post-transcriptional regulation. Plant J **26**: 157–169

Hutchins CJ, Keese P, Visvader JE, Rathjen PD, McInnes JL, Symons RH (1985) Comparison of multimeric plus and minus forms of viroids and virusoids. Plant Mol Biol **4**: 293–304

Hutchins CJ, Rathjen PD, Forster AC, Symons RH (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. Nucleic Acids Res **14**: 3627–3640

Inaba T, Schnell DJ (2008) Protein trafficking to plastids: one theme, many variations. Biochem J **413**: 15–28

Islam W, Noman A, Qasim M, Wang L (2018) Plant Responses to Pathogen Attack: Small RNAs in Focus. Int J Mol Sci **19**: 515

Itaya A, Folimonov A, Matsuda Y, Nelson RS, Ding B (2001) Potato spindle tuber viroid as inducer of RNA silencing in infected tomato. Mol Plant-Microbe Interact **14**: 1332–1334

Itaya A, Matsuda Y, Gonzales RA, Nelson RS, Ding B (2002) Potato spindle tuber viroid Strains of Different Pathogenicity Induces and Suppresses Expression of Common and Unique Genes in Infected Tomato. Mol Plant-Microbe Interact **15**: 990–999

Itaya A, Zhong X, Bundschuh R, Qi Y, Wang Y, Takeda R, Harris AR, Molina C, Nelson RS, Ding B (2007) A Structured Viroid RNA Serves as a Substrate for Dicer-Like Cleavage To Produce Biologically Active Small RNAs but Is Resistant to RNA-Induced Silencing Complex-Mediated Degradation. J Virol **81**: 2980–2994

Jabre I, Reddy ASN, Kalyna M, Chaudhary S, Khokhar W, Byrne LJ, Wilson CM, Syed NH (2019) Does co-transcriptional regulation of alternative splicing mediate plant stress responses? Nucleic Acids Res **47**: 2716–2726

Ji P, Wu W, Chen S, Zheng Y, Zhou L, Zhang J, Cheng H, Yan J, Zhang S, Yang P, and others (2019) Expanded Expression Landscape and Prioritization of Circular RNAs in Mammals. Cell Rep **26**: 3444-3460.e5

Kalantidis K, Denti MA, Tzortzakaki S, Marinou E, Tabler M, Tsagris M (2007) Virp1 is a host protein with a major role in Potato spindle tuber viroid infection in Nicotiana plants. J Virol **81**: 12872–80

Kalinina NO, Makarova S, Makhotenko A, Love AJ, Taliansky M (2018) The Multiple Functions of the Nucleolus in Plant Development, Disease and Stress Responses. Front Plant Sci **9**: 132

Kappagantu M, Bullock JM, Nelson ME, Eastwell KC (2017) Hop stunt viroid: Effect on host (Humulus lupulus) transcriptome and its interactions with hop powdery mildew (Podospheara macularis). Mol Plant-Microbe Interact **30**: 842–851

Karpievitch Y V, Polpitiya AD, Anderson GA, Smith RD, Dabney AR (2010) Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects. Ann Appl Stat **4**: 1797–1823

Katsarou K, Mavrothalassiti E, Dermauw W, Van Leeuwen T, Kalantidis K (2016a) Combined Activity of DCL2 and DCL3 Is Crucial in the Defense against Potato Spindle Tuber Viroid. PLOS Pathog **12**: e1005936

Katsarou K, Wu Y, Zhang R, Bonar N, Morris J, Hedley PE, Bryan GJ, Kalantidis K, Hornyik C (2016b) Insight on Genes Affecting Tuber Development in Potato upon Potato spindle tuber viroid (PSTVd) Infection. PLoS One **11**: e0150711–e0150711

Kerpedjiev P, Hammer S, Hofacker IL (2015) Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. Bioinformatics **31**: 3377–3379

Kim SH, Ryabov E V., Kalinina NO, Rakitina D V., Gillespie T, MacFarlane S, Haupt S, Brown JWS, Taliansky M (2007) Cajal bodies and the nucleolus are required for a plant virus systemic infection. EMBO J **26**: 2169–2179

Kloth KJ, Kormelink R (2020) Defenses against Virus and Vector: A Phloem-Biological Perspective on RTM- and SLI1-Mediated Resistance to Potyviruses and Aphids. Viruses **12**: 129

Kofalvi SA, Marcos JF, Cañizares MC, Pallas V, Candresse T (1997) Hop stunt viroid (HSVd) sequence variants from Prunus species: evidence for recombination between HSVd isolates. J Gen Virol **78**: 3177–3186

Köhler C, Springer N (2017) Plant epigenomics-deciphering the mechanisms of epigenetic inheritance and plasticity in plants. Genome Biol **18**: 132

Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics **27**: 1571–1572

Krupinska K, Blanco NE, Oetke S, Zottini M (2020) Genome communication in plants mediated by organelle–n-ucleus-located proteins. Philos Trans R Soc B Biol Sci **375**: 20190397

Kwan T, Thompson SR (2019) Noncanonical Translation Initiation in Eukaryotes. Cold Spring Harb Perspect Biol **11**: a032672

Lai X, Bazin J, Webb S, Crespi M, Zubieta C, Conn SJ (2018) CircRNAs in plants. *In* J Xiao, red, Adv. Exp. Med. Biol. Springer Singapore, Singapore, bll 329–343

Landry P, Perreault J-P (2005) Identification of a Peach Latent Mosaic Viroid Hairpin Able To Act as a Dicer-Like Substrate. J Virol **79**: 6540–6543

Lecampion C, Floris M, Fantino JR, Robaglia C, Laloi C (2016) An Easy Method for Plant Polysome Profiling. J Vis Exp. doi: 10.3791/54231

Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, and others (2017) Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. Mol Cell **66**: 22-37.e9

Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, and others (2015) Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol **22**: 256–264

Li Z, Zhang Y, Jiang Z, Jin X, Zhang K, Wang X, Han C, Yu J, Li D (2018) Hijacking of the nucleolar protein fibrillarin by TGB1 is required for cell-to-cell movement of Barley stripe mosaic virus. Mol Plant Pathol **19**: 1222–1237

Liang W-C, Wong C-W, Liang P-P, Shi M, Cao Y, Rao S-T, Tsui SK-W, Waye MM-Y, Zhang Q, Fu W-M, and others (2019) Translation of the circular RNA circβ-catenin promotes liver cancer cell growth through activation of the Wnt pathway. Genome Biol **20**: 84

Lin D, O'Callaghan CA (2018) MetClo: methylase-assisted hierarchical DNA assembly using a single type IIS restriction enzyme. Nucleic Acids Res **46**: e113–e113

Lin H-N, Chen C-T, Sung T-Y, Ho S-Y, Hsu W-L (2009) Protein subcellular localization prediction of eukaryotes using a knowledge-based approach. BMC Bioinformatics **10**: S8

Ling J, Luo Z, Liu F, Mao Z, Yang Y, Xie B (2017) Genome-wide analysis of microRNA targeting impacted by SNPs in cucumber genome. BMC Genomics **18**: 275

Liu X, Luo M, Wu K (2012a) Epigenetic interplay of histone modifications and DNA methylation mediated by HDA6. Plant Signal Behav **7**: 633–635

Liu X, Yu CW, Duan J, Luo M, Wang K, Tian G, Cui Y, Wu K (2012b) HDA6 Directly interacts with DNA methyltransferase MET1 and maintains transposable element silencing in Arabidopsis. Plant Physiol **158**: 119–129

Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods **25**: 402–408

López-Carrasco A, Ballesteros C, Sentandreu V, Delgado S, Gago-Zachert S, Flores R, Sanjuán R (2017) Different rates of spontaneous mutation of chloroplastic and nuclear viroids as determined by high-fidelity ultra-deep sequencing. PLoS Pathog **13**: e1006547

López-Carrasco A, Gago-Zachert S, Mileti G, Minoia S, Flores R, Delgado S (2016) The transcription initiation sites of eggplant latent viroid strands map within distinct motifs in their in vivo RNA conformations. RNA Biol **13**: 83–97

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**: 550

Lv DQ, Liu SW, Zhao JH, Zhou BJ, Wang SP, Guo HS, Fang YY (2016) Replication of a pathogenic non-coding RNA increases DNA methylation in plants associated with a bromodomain-containing viroid-binding protein. Sci Rep **6**: 35751

Mandadi KK, Scholthof K-BG (2015) Genome-wide analysis of alternative splicing landscapes modulated during plant-virus interactions in Brachypodium distachyon. Plant Cell **27**: 71–85

Marquez-Molins J, Gomez G, Pallas V (2021) Hop stunt viroid: A polyphagous pathogenic RNA that has shed light on viroid–host interactions. Mol Plant Pathol **22**: 153–162

Marquez-Molins J, Navarro JA, Pallas V, Gomez G (2019) Highly efficient construction of infectious viroid-derived clones. Plant Methods **15**: 87

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**: 10–12

Martín R, Arenas C, Daròs JA, Covarrubias A, Reyes JL, Chua NH (2007) Characterization of small RNAs derived from Citrus exocortis viroid (CEVd) in infected tomato plants. Virology **367**: 135–146

Martínez-Pérez M, Aparicio F, López-Gresa MP, Bellés JM, Sánchez-Navarro JA, Pallás V (2017) Arabidopsis m6A demethylase activity modulates viral infection of a plant virus and the m6A abundance in its genomic RNAs. Proc Natl Acad Sci U S A **114**: 10755–10760

Martínez de Alba AE, Elvira-Matelot E, Vaucheret H (2013) Gene silencing in plants: A diversity of pathways. Biochim Biophys Acta - Gene Regul Mech **1829**: 1300–1308

Martinez de Alba AE, Flores R, Hernandez C (2002) Two Chloroplastic Viroids Induce the Accumulation of Small RNAs Associated with Posttranscriptional Gene Silencing. J Virol **76**: 13094–13096

Martínez F, Marqués J, Salvador ML, Darós JA (2009) Mutational analysis of eggplant latent viroid RNA processing in Chlamydomonas reinhardtii chloroplast. J Gen Virol **90**: 3057–3065

Martinez G, Castellano M, Tortosa M, Pallas V, Gomez G (2014) A pathogenic non-coding RNA induces changes in dynamic DNA methylation of ribosomal RNA genes in host plants. Nucleic Acids Res **42**: 1553–62

Martinez G, Donaire L, Llave C, Pallas V, Gomez G (2010) High-throughput sequencing of Hop stunt viroid-derived small RNAs from cucumber leaves and phloem. Mol Plant Pathol **11**: 347–359

Mastrangelo AM, Marone D, Laidò G, De Leonardis AM, De Vita P (2012) Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. Plant Sci **185**–**186**: 40–49

Matsushita Y, Yanagisawa H, Sano T (2018) Vertical and Horizontal Transmission of Pospiviroids. Viruses **10**: 706

Matzke MA, Kanno T, Matzke AJM (2015) RNA-Directed DNA Methylation: The Evolution of a Complex Epigenetic Pathway in Flowering Plants. Annu Rev Plant Biol **66**: 243–267

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, and others (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. Nature **495**: 333–338

Merchante C, Stepanova AN, Alonso JM (2017) Translation regulation in plants: an interesting past, an exciting present and a promising future. Plant J **90**: 628–653

Minoia S, Carbonell A, Di Serio F, Gisel A, Carrington JC, Navarro B, Flores R (2014) Specific Argonautes Selectively Bind Small RNAs Derived from Potato Spindle Tuber Viroid and Attenuate Viroid Accumulation In Vivo. J Virol **88**: 11933–11945

Minoia S, Navarro B, Delgado S, Di Serio F, Flores R (2015) Viroid RNA turnover: characterization of the subgenomic RNAs of potato spindle tuber viroid accumulating in infected tissues provides insights into decay pathways operating in vivo. Nucleic Acids Res **43**: 2313–25

Molina-Serrano D, Suay L, Salvador ML, Flores R, Daròs J-A (2007) Processing of RNAs of the family Avsunviroidae in Chlamydomonas reinhardtii chloroplasts. J Virol **81**: 4363–6

Mühlbach HP, Sänger HL (1979) Viroid replication is inhibited by α-amanitin. Nature **278**: 185–188

Nair R, Rost B (2003) Better prediction of sub-cellular localization by combining evolutionary and structural information. Proteins Struct Funct Bioinforma **53**: 917–930

Navarro B, Flores R (1997) Chrysanthemum chlorotic mottle viroid: Unusual structural properties of a subgroup of self-cleaving viroids with hammerhead ribozymes. Proc Natl Acad Sci U S A **94**: 11262–11267

Navarro B, Flores R, Di Serio F (2021a) Advances in Viroid-Host Interactions. Annu Rev Virol **8**: 305–325

Navarro B, Gisel A, Rodio ME, Delgado S, Flores R, Di Serio F (2012a) Viroids: How to infect a host and cause disease without encoding proteins. Biochimie **94**: 1474–1480

Navarro B, Gisel A, Rodio ME, Delgado S, Flores R, Di Serio F (2012b) Small RNAs containing the pathogenic determinant of a chloroplast- replicating viroid guide the degradation of a host mRNA as predicted by RNA silencing. Plant J **70**: 991–1003

Navarro B, Gisel A, Serra P, Chiumenti M, Di Serio F, Flores R (2021b) Degradome Analysis of Tomato and Nicotiana benthamiana Plants Infected with Potato Spindle Tuber Viroid. Int J Mol Sci **22**: 3725

Navarro B, Pantaleo V, Gisel A, Moxon S, Dalmary T, Bisztray G, Di Serio F, Burgyán J (2009) Deep sequencing of viroid-derived small RNAs from grapevine provides new insights on the role of RNA silencing in plant-viroid interaction. PLoS One **4**: e7686

Navarro JA, Vera A, Flores R (2000) A chloroplastic RNA polymerase resistant to tagetitoxin is involved in replication of Avocado sunblotch viroid. Virology **268**: 218–225

Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, Kinzler KW, Vogelstein B (1991) Scrambled exons. Cell **64**: 607–613

Nohales M-Á, Flores R, Daròs J-A (2012a) Viroid RNA redirects host DNA ligase 1 to act as an RNA ligase. Proc Natl Acad Sci **109**: 13805–13810

Nohales M-A, Molina-Serrano D, Flores R, Daros J-A (2012b) Involvement of the Chloroplastic Isoform of tRNA Ligase in the Replication of Viroids Belonging to the Family Avsunviroidae. J Virol **86**: 8269–8276

Owens RA, Blackburn M, Ding B (2001) Possible involvement of the phloem lectin in long-distance viroid movement. Mol Plant-Microbe Interact **14**: 905–909

Owens RA, Gomez G, Lison P, Conejero V (2017) Changes in the Host Proteome and Transcriptome Induced by Viroid Infection. In: Viroids and Satellites. Chapter 10. Edited by A. Hadidi, R. Flores, J. W. Randles and P. Palukaitis. Academic Press, Elsevier. 754 pp. pp. 105–114 ISBN: 978-0-12-801498-1.

Owens RA, Hammond RW (2009) Viroid pathogenicity: One process, many faces. Viruses **1**: 298–316

Owens RA, Tech KB, Shao JY, Sano T, Baker CJ (2012) Global analysis of tomato gene expression during potato spindle tuber viroid infection reveals a complex array of changes affecting hormone signaling. Mol Plant-Microbe Interact **25**: 582–598

Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez R, Warthmann N, Allen E, Dezulian T, Huson D, and others (2007) Sequence and Expression Differences Underlie Functional Specialization of Arabidopsis MicroRNAs miR159 and miR319. Dev Cell **13**: 115–125

Pallas V, Gómez G (2013) Phloem RNA-binding proteins as potential components of the long-distance RNA transport system. Front Plant Sci **4**: 1–6

Pallás V, Gómez G (2017) Viroid Movement. In: Viroids and Satellites. Chapter 8. Edited by A. Hadidi, R. Flores, J. W. Randles and P. Palukaitis. Academic Press, Elsevier. 754 pp. pp. 83-91 ISBN: 978-0-12-801498-1.

Pallas V, Navarro A, Flores R (1987) Isolation of a viroid-like RNA from hop different from hop stunt viroid. J Gen Virol **68**: 3201–3205

Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, and others (2017) Translation of CircRNAs. Mol Cell **66**: 9-21.e7

Papaefthimiou I, Hamilton A, Denti M, Baulcombe D, Tsagris M, Tabler M (2001) Replicating potato spindle tuber viroid RNA is accompanied by short RNA fragments that are characteristic of post-transcriptional gene silencing. Nucleic Acids Res **29**: 2395–400

Patop IL, Wüst S, Kadener S (2019) Past, present, and future of circRNAs. EMBO J **38**: e100836

Pignatta D, Novitzky K, Satyaki PRV, Gehring M (2018) A variably imprinted epiallele impacts seed development. PLOS Genet **14**: e1007469

Prats AC, David F, Diallo LH, Roussel E, Tatin F, Garmy-Susini B, Lacazette E (2020) Circular rna, the key for translation. Int J Mol Sci **21**: 1–18

Pumplin N, Voinnet O (2013) RNA silencing suppression by plant pathogens: Defence, counter-defence and counter-counter-defence. Nat Rev Microbiol **11**: 745–760

Qi Y, Ding B (2003) Differential Subnuclear Localization of RNA Strands of Opposite Polarity Derived from an Autonomously Replicating Viroid. Plant Cell **15**: 2566–2577

Qi Y, Ding B (2002) Replication of Potato spindle tuber viroid in cultured cells of tobacco and Nicotiana benthamiana: the role of specific nucleotides in determining replication levels for host adaptation. Virology **302**: 445–456

Qi Y, Pélissier T, Itaya A, Hunt E, Wassenegger M, Ding B (2004) Direct role of a viroid RNA motif in mediating directional RNA trafficking accross a specific cellular boundary. Plant Cell **16**: 1741–1752

Qu F, Ye X, Hou G, Sato S, Clemente TE, Morris TJ (2005) RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in Nicotiana benthamiana. J Virol **79**: 15209–17

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics **26**: 841–842

Raja P, Sanville BC, Buchmann RC, Bisaro DM (2008) Viral Genome Methylation as an Epigenetic Defense against Geminiviruses. J Virol **82**: 8997–9007

Ramesh S V., Yogindran S, Gnanasekaran P, Chakraborty S, Winter S, Pappu HR (2021) Virus and Viroid-Derived Small RNAs as Modulators of Host Gene Expression: Molecular Insights Into Pathogenesis. Front Microbiol **11**: 3170

Rathinasabapathi B, Burnet M, Russell BL, Gage DA, Liao P-C, Nye GJ, Scott P, Golbeck JH, Hanson AD (1997) Choline monooxygenase, an unusual iron-sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants: Prosthetic group characterization and cDNA cloning. Proc Natl Acad Sci **94**: 3454 LP – 3458

Richter S, Lamppa GK (2002) Determinants for removal and degradation of transit peptides of chloroplast precursor proteins. J Biol Chem **277**: 43888–43894

Riesner D, Henco K, Rokohl U, Klotz G, Kleinschmidt AK, Domdey H, Jank P, Gross HJ, Sänger HL (1979) Structure and structure formation of viroids. J Mol Biol **133**: 85–115

Rizza S, Conesa A, Juarez J, Catara A, Navarro L, Duran-Vila N, Ancillo G (2012) Microarray analysis of etrog citron (citrus medica l.) reveals changes in chloroplast, cell wall, peroxidase and symporter activities in response to viroid infection. Mol Plant Pathol **13**: 852–864

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**: 139

Rodio ME, Delgado S, De Stradis A, Gómez MD, Flores R, Di Serio F (2007) A viroid RNA with a specific structural motif inhibits chloroplast development. Plant Cell **19**: 3610–3626

Rodríguez-Negrete E, Lozano-Durán R, Piedra-Aguilera A, Cruzado L, Bejarano ER, Castillo AG (2013) Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytol **199**: 464–475

Sang HK, MacFarlane S, Kalinina NO, Rakitina D V., Ryabov E V., Gillespie T, Haupt S, Brown JWS, Taliansky M (2007) Interaction of a plant virus-encoded protein with the major nucleolar protein fibrillarin is required for systemic virus infection. Proc Natl Acad Sci U S A **104**: 11115–11120

Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK (1976) Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. Proc Natl Acad Sci **73**: 3852–3856

Sanjuán R, Daròs J-A (2007) One-step site-directed mutagenesis of viroid dimeric cDNA. J Virol Methods **145**: 71–75

Sano T (2021) Progress in 50 years of viroid research-Molecular structure, pathogenicity, and host adaptation. Proc Jpn Acad Ser B Phys Biol Sci **97**: 371–401

Sano T, Matsuura Y (2004) Accumulation of short interfering RNAs characteristic of RNA silencing precedes recovery of tomato plants from severe symptoms of Potato spindle tuber viroid infection. J Gen Plant Pathol **70**: 50–53

Sanz-Carbonell A, Marques MC, Bustamante A, Fares MA, Rodrigo G, Gomez G (2019) Inferring the regulatory network of the miRNA-mediated response to biotic and abiotic stress in melon. BMC Plant Biol **19**: 78

Sanz-Carbonell A, Marques MC, Martinez G, Gomez G (2020) Dynamic architecture and regulatory implications of the miRNA network underlying the response to stress in melon. RNA Biol **17**: 292–308

Sasaki M, Shikata E (1977) On Some Properties of Hop Stunt Disease Agent, a Viroid. Proc Japan Acad Ser B Phys Biol Sci **53**: 109–112

Sasaki T, Lee T, Liao W-W, Naumann U, Liao J-L, Eun C, Huang Y-Y, Fu JL, Chen P-Y, Meyers BC, and others (2014) Distinct and concurrent pathways of Pol II- and Pol IV-dependent siRNA biogenesis at a repetitive trans -silencer locus in Arabidopsis thaliana. Plant J **79**: 127–138

Schwind N, Zwiebel M, Itaya A, Ding B, Wang MB, Krczal G, Wassenegger M (2009) RNAi-mediated resistance to Potato spindle tuber viroid in transgenic tomato expressing a viroid hairpin RNA construct. Mol Plant Pathol **10**: 459–469

Serra P, Messmer A, Sanderson D, James D, Flores R (2018) Apple hammerhead viroid-like RNA is a bona fide viroid: Autonomous replication and structural features support its inclusion as a new member in the genus Pelamoviroid. Virus Res **249**: 8–15

Shang X, Cao Y, Ma L (2017) Alternative Splicing in Plant Genes: A Means of Regulating the Environmental Fitness of Plants. Int J Mol Sci **18**: 432

Shi T, Gaffrey MJ, Fillmore TL, Nicora CD, Yi L, Zhang P, Shukla AK, Wiley HS, Rodland KD, Liu T, and others (2018) Facile carrier-assisted targeted mass spectrometric approach for proteomic analysis of low numbers of mammalian cells. Commun Biol **1**: 103

Silmon De Monerri NC, Kim K (2014) Pathogens hijack the epigenome: A new twist on host-pathogen interactions. Am J Pathol **184**: 897–911

Simpson GG, Laurie RE, Dijkwel PP, Quesada V, Stockwell PA, Dean C, Macknight RC (2010) Noncanonical translation initiation of the Arabidopsis flowering time and alternative polyadenylation regulator FCA. Plant Cell **22**: 3764–3777

Škorić D (2017) Viroid Biology. In: Viroids and Satellites. Chapter 5. Edited by A. Hadidi, R. Flores, J. W. Randles and P. Palukaitis. Academic Press, Elsevier. 754 pp. pp. 53–61 ISBN: 978-0-12-801498-1.

Slotkin RK (2016) Plant epigenetics: from genotype to phenotype and back again. Genome Biol **17**: 57

Spiesmacher E, Mühlbach HP, Schnölzer M, Haas B, Sänger HL (1983) Oligomeric forms of potato spindle tuber viroid (PSTV) and of its complementary RNA are present in nuclei isolated from viroid-infected potato cells. Biosci Rep **3**: 767–774

Štajner N, Radišek S, Mishra AK, Nath VS, Matoušek J, Jakše J (2019) Evaluation of Disease Severity and Global Transcriptome Response Induced by Citrus bark cracking viroid, Hop latent viroid, and Their Co-Infection in Hop (Humulus lupulus L.). Int J Mol Sci **20**: 3154

Steger G, Riesner D (2018) Viroid research and its significance for RNA technology and basic biochemistry. Nucleic Acids Res **46**: 10563–10576

Sweeney BA, Petrov AI, Burkov B, Finn RD, Bateman A, Szymanski M, Karlowski WM, Gorodkin J, Seemann SE, Cannone JJ, and others (2019) RNAcentral: A hub of information for non-coding RNA sequences. Nucleic Acids Res **47**: D221–D229

Symons RH (1981) Avocado sunblotch viroid: primary sequence and proposed secondary structure. Nucleic Acids Res **9**: 6527–37

Szittya G, Silhavy D, Molnár A, Havelda Z, Lovas A, Lakatos L, Bánfalvi Z, Burgyán J (2003) Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. EMBO J **22**: 633–40

Tabler M, Sänger HL (1985) Infectivity studies on different potato spindle tuber viroid (PSTV) RNAs synthesized in vitro with the SP6 transcription system. EMBO J **4**: 2191–9

Takeda R, Petrov AI, Leontis NB, Ding B (2011) A three-dimensional RNA motif in Potato spindle tuber viroid mediates trafficking from palisade mesophyll to spongy mesophyll in Nicotiana benthamiana. Plant Cell **23**: 258–72

Takino H, Kitajima S, Hirano S, Oka M, Matsuura T, Ikeda Y, Kojima M, Takebayashi Y, Sakakibara H, Mino M (2019) Global transcriptome analyses reveal that infection with chrysanthemum stunt viroid (CSVd) affects gene expression profile of chrysanthemum plants, but the genes involved in plant hormone metabolism and signaling may not be silencing target of CSVd-siRNAs. Plant Gene **18**: 100181

Taliansky ME, Brown JWS, Rajamäki ML, Valkonen JPT, Kalinina NO (2010) Involvement of the plant nucleolus in virus and viroid infections: parallels with animal pathosystems. Adv Virus Res **77**: 119–158

Tamura K, Hara-Nishimura I (2014) Functional insights of nucleocytoplasmic transport in plants. Front Plant Sci **5**: 1–10

Tan J, Zhou Z, Niu Y, Sun X, Deng Z (2017) Identification and Functional Characterization of Tomato CircRNAs Derived from Genes Involved in Fruit Pigment Accumulation. Sci Rep **7**: 8594

Thibaut O, Claude B (2018) Innate Immunity Activation and RNAi Interplay in Citrus Exocortis Viroid—Tomato Pathosystem. Viruses **10**: 587

Torchetti EM, Pegoraro M, Navarro B, Catoni M, Di Serio F, Noris E (2016) A nuclear-replicating viroid antagonizes infectivity and accumulation of a geminivirus by upregulating methylation-related genes and inducing hypermethylation of viral DNA. Sci Rep **6**: 35101

Tsuda K, Somssich IE (2015) Transcriptional networks in plant immunity. New Phytol **206**: 932–947

Tsushima D, Adkar-Purushothama CR, Taneda A, Sano T (2015) Changes in relative expression levels of viroid-specific small RNAs and microRNAs in tomato plants infected with severe and mild symptom-inducing isolates of Potato spindle tuber viroid. J Gen Plant Pathol **81**: 49–62

Ueki S, Citovsky V (2001) RNA commutes to work: Regulation of plant gene expression by systemically transported RNA molecules. BioEssays **23**: 1087–1090

Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci **13**: 350–358

Verhoeven JTJ, Meekes ETM, Roenhorst JW, Flores R, Serra P (2013) Dahlia latent viroid: A recombinant new species of the family Pospiviroidae posing intriguing questions about its origin and classification. J Gen Virol **94**: 711–719

Visvader JE, Forster AC, Symons RH (1985) Infectivity and in vitro mutagenesis of monomeric cDNA clones of citrus exocortis viroid indicates the site of processing of viroid precursors. Nucleic Acids Res **13**: 5843–56

Wang C, Wang C, Zou J, Yang Y, Li Z, Zhu S (2019a) Epigenetics in the plant–virus interaction. Plant Cell Rep **38**: 1031–1038

Wang MB, Bian XY, Wu LM, Liu LX, Smith NA, Isenegger D, Wu RM, Masuta C, Vance VB, Watson JM, and others (2004) On the role of RNA silencing in the pathogenicity and evolution of viroids and viral satellites. Proc Natl Acad Sci U S A **101**: 3275–3280

Wang X, Li N, Li W, Gao X, Cha M, Qin L, Liu L (2020) Advances in Transcriptomics in the Response to Stress in Plants. Glob Med Genet **7**: 30–34

Wang Y (2021) Current view and perspectives in viroid replication. Curr Opin Virol **47**: 32–37

Wang Y, Shibuya M, Taneda A, Kurauchi T, Senda M, Owens RA, Sano T (2011) Accumulation of Potato spindle tuber viroid-specific small RNAs is accompanied by specific changes in gene expression in two tomato cultivars. Virology **413**: 72–83

Wang Y, Wu J, Qiu Y, Atta S, Zhou C, Cao M (2019b) Global transcriptomic analysis reveals insights into the response of 'etrog' citron (Citrus medica L.) to Citrus Exocortis viroid infection. Viruses **11**: 453

Wang Y, Yang M, Wei S, Qin F, Zhao H, Suo B (2017a) Identification of Circular RNAs and Their Targets in Leaves of Triticum aestivum L. under Dehydration Stress. Front Plant Sci **7**: 2024

Wang Z, Liu Y, Li D, Li L, Zhang Q, Wang S, Huang H (2017b) Identification of Circular RNAs in Kiwifruit and Their Species-Specific Response to Bacterial Canker Pathogen Invasion. Front Plant Sci **8**: 413

Warrilow D, Symons RH (1999) Citrus exocortis viroid RNA is associated with the largest subunit of RNA polymerase II in tomato in vivo. Arch Virol **144**: 2367–2375

Wassenegger M, Dalakouras A (2021) Viroids as a Tool to Study RNA-Directed DNA Methylation in Plants. Cells **10**: 1187

Wassenegger M, Heimes S, Riedel L, Sänger HL (1994) RNA-directed de novo methylation of genomic sequences in plants. Cell **76**: 567–76

Waters AJ, Makarevitch I, Noshay J, Burghardt LT, Hirsch CN, Hirsch CD, Springer NM (2017) Natural variation for gene expression responses to abiotic stress in maize. Plant J **89**: 706–717

Wei S, Bian R, Andika IB, Niu E, Liu Q, Kondo H, Yang L, Zhou H, Pang T, Lian Z, and others (2019) Symptomatic plant viroid infections in phytopathogenic fungi. Proc Natl Acad Sci **116**: 13042 LP – 13050

Wendte JM, Pikaard CS (2017) The RNAs of RNA-directed DNA methylation. Biochim Biophys Acta - Gene Regul Mech **1860**: 140–148

Wernersson R (2006) Virtual Ribosome--a comprehensive DNA translation tool with support for integration of sequence feature annotation. Nucleic Acids Res **34**: W385–W388

Wesselhoeft RA, Kowalski PS, Anderson DG (2018) Engineering circular RNA for potent and stable translation in eukaryotic cells. Nat Commun **9**: 2629

Whitham SA, Anderberg RJ, Chisholm ST, Carrington JC (2000) Arabidopsis RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein. Plant Cell **12**: 569–582

Wickham H (2011) ggplot2. Wiley Interdiscip Rev Comput Stat **3**: 180–185

Wilkins O, Hafemeister C, Plessis A, Holloway-Phillips M-M, Pham GM, Nicotra AB, Gregorio GB, Jagadish SVK, Septiningsih EM, Bonneau R, and others (2016) EGRINs (Environmental Gene Regulatory Influence Networks) in Rice That Function in the Response to Water Deficit, High Temperature, and Agricultural Environments. Plant Cell **28**: 2365–2384

Williams BP, Pignatta D, Henikoff S, Gehring M (2015) Methylation-Sensitive Expression of a DNA Demethylase Gene Serves As an Epigenetic Rheostat. PLOS Genet **11**: e1005142

Wu TD, Watanabe CK (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics **21**: 1859–1875

Xia C, Li S, Hou W, Fan Z, Xiao H, Lu M, Sano T, Zhang Z (2017) Global transcriptomic changes induced by infection of cucumber (Cucumis sativus L.) with mild and severe variants of hop stunt viroid. Front Microbiol **8**: 1–16

Xu X, Wang K, Pan J, Chen X (2019) Small RNA sequencing identifies cucumber miRNA roles in waterlogging-triggered adventitious root primordia formation. Mol Biol Rep **46**: 6381–6389

Xu X, Zhong C, Tan M, Song Y, Qi X, Xu Q, Chen X (2020) Identification of MicroRNAs and Their Targets That Respond to Powdery Mildew Infection in Cucumber by Small RNA and Degradome Sequencing. Front Genet **11**: 246

Yaish MW (2017) Editorial: Epigenetic Modifications Associated with Abiotic and Biotic Stresses in Plants: An Implication for Understanding Plant Evolution. Front Plant Sci **8**: 1983

Yan J, Gu Y, Jia X, Kang W, Pan S, Tang X, Chen X, Tang G (2012) Effective Small RNA Destruction by the Expression of a Short Tandem Target Mimic in Arabidopsis. Plant Cell **24**: 415–427

Yang L-P, Fang Y-Y, An C-P, Dong L, Zhang Z-H, Chen H, Xie Q, Guo H-S (2013) C2-mediated decrease in DNA methylation, accumulation of siRNAs, and increase in expression for genes involved in defense pathways in plants infected with beet severe curly top virus. Plant J **73**: 910–917

Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen L-L, Wang Y, and others (2017) Extensive translation of circular RNAs driven by N(6)-methyladenosine. Cell Res **27**: 626–641

Yang Y, Hsu PJ, Chen Y-S, Yang Y-G (2018) Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res **28**: 616–624

Ye C-Y, Chen L, Liu C, Zhu Q-H, Fan L (2015) Widespread noncoding circular RNAs in plants. New Phytol **208**: 88–95

Zhai J, Arikit S, Simon SA, Kingham BF, Meyers BC (2014) Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. Methods **67**: 84–90

Zhang H, Lang Z, Zhu JK (2018a) Dynamics and function of DNA methylation in plants. Nat Rev Mol Cell Biol **19**: 489–506

Zhang M, Huang N, Yang X, Luo J, Yan S, Xiao F, Chen W, Gao X, Zhao K, Zhou H, and others (2018b) A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. Oncogene **37**: 1805–1814

Zhang M, Zhao K, Xu X, Yang Y, Yan S, Wei P, Liu H, Xu J, Xiao F, Zhou H, and others (2018c) A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. Nat Commun **9**: 4475

Zhang X, Bernatavichute Y V, Cokus S, Pellegrini M, Jacobsen SE (2009) Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. Genome Biol **10**: R62

Zhang X, Clarenz O, Cokus S, Bernatavichute Y V, Pellegrini M, Goodrich J, Jacobsen SE (2007) Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis. PLoS Biol **5**: e129

Zhang Z, Xia C, Matsuda T, Taneda A, Murosaki F, Hou W, Owens RA, Li S, Sano T (2020) Effects of Host-Adaptive Mutations on Hop Stunt Viroid Pathogenicity and Small RNA Biogenesis. Int J Mol Sci **21**: 7383

Zhao W, Cheng Y, Zhang C, You Q, Shen X, Guo W, Jiao Y (2017) Genome-wide identification and characterization of circular RNAs by high throughput sequencing in soybean. Sci Rep **7**: 5636

Zhao W, Chu S, Jiao Y (2019) Present scenario of circular RNAs (circRNAs) in plants. Front Plant Sci. doi: 10.3389/fpls.2019.00379

Zhao Y, Owens RA, Hammond RW (2001) Use of a vector based on Potato virus X in a whole plant assay to demonstrate nuclear targeting of potato spindle tuber viroid. J Gen Virol **82**: 1491–1497

Zheng L, Du Z, Lin C, Mao Q, Wu K, Wu J, Wei T, Wu Z, Xie L (2015) Rice stripe tenuivirus p2 may recruit or manipulate nucleolar functions through an interaction with fibrillarin to promote virus systemic movement. Mol Plant Pathol **16**: 921–930

Zheng X, Chen L, Zhou Y, Wang Q, Zheng Z, Xu B, Wu C, Zhou Q, Hu W, Wu C, and others (2019a) A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. Mol Cancer **18**: 47

Zheng Y, Wang Y, Ding B, Fei Z (2017) Comprehensive Transcriptome Analyses Reveal that Potato Spindle Tuber Viroid Triggers Genome-Wide Changes in Alternative Splicing, Inducible trans-Acting Activity of Phased Secondary Small Interfering RNAs, and Immune Responses. J Virol **91**: e00247-17

Zheng Y, Wu S, Bai Y, Sun H, Jiao C, Guo S, Zhao K, Blanca J, Zhang Z, Huang S, and others (2019b) Cucurbit Genomics Database (CuGenDB): a central portal for comparative and functional genomics of cucurbit crops. Nucleic Acids Res **47**: D1128–D1136

Zhou J, Wang X, He K, Charron JBF, Elling AA, Deng XW (2010) Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in arabidopsis reveals correlation between multiple histone marks and gene expression. Plant Mol Biol **72**: 585–595

Zhu Q-H, Shan W-X, Ayliffe MA, Wang M-B (2016) Epigenetic Mechanisms: An Emerging Player in Plant-Microbe Interactions. Mol Plant-Microbe Interact **29**: 187–196

Zhu Y, Qi Y, Xun Y, Owens R, Ding B (2002) Movement of potato spindle tuber viroid reveals regulatory points of phloem-mediated RNA traffic. Plant Physiol **130**: 138–146

Zhu YX, Jia JH, Yang L, Xia YC, Zhang HL, Jia JB, Zhou R, Nie PY, Yin JL, Ma DF, and others (2019) Identification of cucumber circular RNAs responsive to salt stress. BMC Plant Biol **19**: 164

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res **31**: 3406–15