Diversity of the mechanisms related to viral RNA uridylation in plants

Anne-Caroline Joly

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Institut de biologie moléculaire des plantes – CNRS – UPR 2357

THÈSE présentée par :

Anne Caroline JOLY

Soutenue le : **7 octobre 2022**

Pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

Diversity of the mechanisms related to viral RNA uridylation in plants

Directrice de thèse :	Dr. Hélène ZUBER, Université de Strasbourg
Co-directeur de thèse :	Dr. Dominique GAGLIARDI, Université de Strasbourg
Rapporteur Externe :	Prof. Štěpánka Vaňáčová, Masaryk University
Rapporteur Externe :	Dr. Stéphane BLANC, INRA Montpellier
Examinateur Externe :	Prof. Peter REDDER, Université Toulouse III
Examinateur Interne :	Dr. Sébastien PFEFFER, Université de Strasbourg

A la mémoire de Henri Joly

Acknowledgements

First, I would like to thank Prof. Štěpánka Vaňáčová and Dr. Stéphane Blanc for accepting to evaluate this manuscript, as well as Prof. Peter Redder and Dr. Sébastien Pfeffer that examined this manuscript as well.

Tant de personnes m'ont aidée au cours de ma thèse, mais également pendant ma formation universitaire, qu'il me serait impossible de nommer tout le monde ici. Ces remerciements ne sont donc pas exhaustifs.

Je tiens d'abord à remercier tout particulièrement Hélène et Gag, qui m'ont encadrée depuis mon premier stage de Master 1, puis pendant mon stage de Master 2 et ma thèse.
Hélène, merci pour ta patience, ta disponibilité et ton aide pendant ces quatre ans et demi. Tu m'as significativement donné goût pour la bioinformatique (et les stats !). Merci d'avoir toujours répondu à mes nombreuses (et naïves) questions et de m'avoir encouragée.
Gag, je te suis très reconnaissante de merci de m'avoir accueillie dans l'équipe en stage de Master 1. Je te remercie pour ton implication et tes conseils, tant sur le plan scientifique qu'humain. Merci de m'avoir fait confiance dans le cadre de ce nouveau projet innovant et passionnant sur les ARN viraux.

Ensuite, je remercie le reste de l'équipe qui contribue à une bonne ambiance de travail :
Damien pour son aide sur le projet, ses bonnes blagues et son calme olympien,
Heike, thank you for the helpful discussions and the smart technical tips,
Aude, Quentin et Pietro, merci pour votre soutien, ce fut un plaisir de partager votre bureau. Aude, merci pour ton écoute, tu es une très belle personne et je te souhaite le meilleur pour ta future carrière. Quentin, merci pour ton aide, force et honneur pour la suite, je n'oublierai pas notre rencontre au labo puis à la Laiterie. Pietro, merci d'avoir été un bon compagnon de commérages ;)
Elodie, merci pour ta gentillesse et ton aide dans la sélection des lignées.
J'ai aussi une pensée pour Caroline De Almeida et Marlène Schiaffini, anciennes doctorantes dans l'équipe qui m'ont beaucoup aidée à mon arrivée et jusqu'au début de ma thèse. Je remercie également les deux stagiaires que j'ai eu l'occasion d'encadrer, Samuel Probst pour son aide dans la sélection de lignées transgéniques et Lucie Labeauvie.

Impossible de ne pas remercier Shah pour sa gentillesse et son écoute, ainsi que son implication et sa persévérance dans les manips de 3'RACE-seq sur les ARN viraux. Merci à

Manue et Jim pour vos réponses à mes questions souvent naïves et pour les discussions intéressantes.
Un grand merci à Sandrine et Malek de la plateforme de séquençage de l'IBMP. Je remercie aussi tous les membres de l'IBMP rencontré.e.s au cours de ces dernières années, qui m'ont apporté au niveau scientifique et personnel. Aux gens en salle de rédaction, bon courage pour la dernière ligne droite !
Merci ensuite à tout le personnel de l'IBMP des services techniques et administratifs pour leur aide, notamment le service de production de plantes, sans qui la culture d'Arabidopsis et Bentha serait bien plus compliquée.

Je tiens aussi à remercier les chercheur.e.s qui ont contribué à la collection de virus analysés en 3'RACE-seq : Salah Bouzoubaah, Véronique Brault, Marc Fuchs, David Gilmer, Anthony Gobert, Manfred Heinlein, Jean-Sébastien Reynard, Christophe Ritzenthaler, Corinne Schmitt-Keichinger, Claire Villeroy and Véronique Ziegler-Graff. Je remercie particulièrement David Gilmer pour les conseils sur le projet. Merci à Anthony pour son astuce afin d'optimiser la ligation et à Jérôme pour les macros ImageJ.

J'ai également une pensée pour Joern Pütz, que je remercie pour m'avoir donné la possibilité de faire partie du cursus franco-allemand en licence, ce qui m'a donné beaucoup d'opportunités professionnelles mais aussi beaucoup apporté sur le plan personnel.
Je remercie aussi Adrien Franchet, qui m'a encadrée lors de mon premier stage en laboratoire à l'IBMC, pour ses encouragements à persévérer dans le monde de la recherche.

Je remercie mes ami.e.s pour leur soutien, sans qui ces dernières années auraient été beaucoup plus difficiles : Alison, Marie, Julie, Fanny, Camille, Jerem, Nico, Vincent, Valentin, Léna, Célia, l'ensemble de la team Odyssee. Alison, merci d'avoir toujours cru en moi et d'avoir toujours été là pour m'encourager, pour tous ces beaux moments passés ensemble et ceux à venir.
Enfin, mon plus grand merci ira à François et à ma famille. François, merci pour ta présence à toute épreuve et le réconfort dans les moments compliqués. Merci de m'avoir appris à respirer et à être plus patiente. Merci à ma famille pour leur soutien sans faille : Aurel, qui a toujours été là quand ça n'allait pas, Maman et Papa merci pour votre implication, vos encouragements et pour m'avoir appris à donner le meilleur de moi-même. Un grand merci aussi à Mamée (qui se réjouit, je crois, que je sois passée de la drosophile à Arabidopsis) et à Pierre et Mounh. Je vous suis grandement reconnaissante pour votre soutien tout au long de ma scolarité, en particulier durant les trois dernières années.

To Charlotte de Witte, Ghost, Talamasca, Astrix, Gioli and Assia, Billx, Nils Frahm, D-Block, Noisecontrollers, Enrico Sangiuliano, 99999999, Liquid Soul, Electric Universe, Le Wanski, Stella Bossi, A*S*Y*S, Paul van Dyk, Space 92, Alignement, Indochine, Pink Floyd, Caesaria, Black Sabbath, Arctic Monkeys, Angèle, Dikanda, Infected Mushroom, Raja Ram, Pawlowski, Ovnimoon, Indira Paganotto, Kalki, AIROD, Trym, VCL, Scorpions, Avatar, Ludovico Einaudi, Fabrizio Patterlini, Hans Zimmer, John Williams, Kabylie Minogue, Acid Arab, Howard Shore, Mauro Picotto, Billx, Dr Peacock, Paul Kalkbrenner and others whose music helped me through the writing of this manuscript and the whole thesis.

Table of contents

List of figures

List of tables

Abreviations

3'CITE	3' Cap-independent translation enhancer
3’RACE-seq	3’ Rapid amplification of cDNA ends sequencing
cDNA	Complementary DNA
cRNA	Complementary RNA
Ct	Cycle threshold
DCL	DICER-LIKE
DI RNA	Defective interfering RNA
DRS	Direct RNA sequencing
ENE	Element for nuclear expression
ETI	Effector-triggered immunity
gRNA	Genomic RNA
HR	Hypersensitive response
ICTV	International committee on taxonomy of viruses
IRES	Internal ribosome entry site
MAMP	Microbe-associated molecular pattern
miRNA	Micro RNA
mRNA	Messenger RNA
NEXT complex	Nuclear exosome targeting complex
NMD	Nonsense-mediated decay
ORF	Open reading frame
PABP	Poly(A) binding protein
PABPC	Cytosolic poly(A) binding protein
PABPN	Nuclear poly(A) binding protein
PAMP	Pathogen-associated molecular pattern
PAP	Poly(A) polymerase
PAXT connection	Poly(A) exosome targeting connection
PEG	Polyethylene glycol
PIPO	Pretty interesting *Potyvirus* ORF
PRR	Pattern recognition receptor
PTI	Pathogen-triggered immunity
RdRp	RNA-dependent RNA polymerase
RRM	RNA recognition motif
satRNA	Satellite RNA
sgRNA	Subgenomic RNA
SHAPE	Selective 2’-hydroxyl acylation analyzed by primer extension
siRNA	Small interfering RNA
SMRT® sequencing	Single molecule real-time® sequencing
TGBp	Triple gene block protein
TLS	tRNA-like structure

TNTase	Terminal nucleotidyl transferase
TRAMP complex	Trf4/5-Air1/2-Mtr4 polyadenylation complex
TRUMP complex	Terminal RNA uridylation-mediated processing complex
TSS	T-shaped structure
TUTase	Terminal uridylyl transferase
USR	Unstructured region
VPg	Viral protein genome-linked
VSR	Viral silencing suppressor

Viruses mentioned in this study

To facilitate the reading, viral species are abbreviated in the text by their common acronym.

AlRV	*Alternaria longipes dsRNA virus*
AlMV	*Alfalfa mosaic virus*
ArMV	*Arabis mosaic virus*
BaMV	*Bamboo mosaic virus*
BBWV-1	*Broad bean mottle virus 1*
BMV	*Brome mosaic virus*
BNYVV	*Beet necrotic yellow vein virus*
CLRV	*Cherry leafroll virus*
CMV	*Cucumber mosaic virus*
CNDV	*Carrot necrotic dieback virus*
CNV	*Cucumber necrosis virus*
CPMV	*Cowpea mosaic virus*
DuMV	*Dulcamara mottle virus*
FgHV2	*Fusarium graminearum hypovirus 2*
GFLV	*Grapevine fanleaf virus*
GLRaV-1	*Grapevine leafroll-associated virus 1*
GLRaV-2	*Grapevine leafroll-associated virus 2*
GPGV	*Grapevine pinot gris virus*
GRGV	*Grapevine red globe virus*
GRSPaV	*Grapevine rupestris stem pitting-associated virus*
GVB	*Grapevine virus B*
HCV	*Hepatitis C virus*
IAV	*Influenza A virus*
NeRNV	*Nemesia ring necrosis virus*
ORMV	*Oilseed rape mosaic virus*
ORSV	*Odontoglossum ringspot virus*
OrV	*Orsay virus*
PCV	*Peanut clump virus*
PEDV	*Porcine epidemic diarrhea virus*
PlMAV	*Plantago asiatica mosaic virus*

PPV	*Plum pox virus*
PRRSV	*Porcine respiratory and reproductive syndrome virus*
PVX	*Potato virus X*
PVY	*Potato virus Y*
RDV	*Rice dwarf virus*
RpRSV	*Raspberry ringspot virus*
SFV	*Semliki Forest virus*
SoMV	*Sowbane mosaic virus*
SLRSV	*Strawberry latent ringspot virus*
TBRV	*Tomato black ring virus*
TBSV	*Tomato bushy stunt virus*
TCV	*Turnip crinkle virus*
TEV	*Tobacco etch virus*
TMV	*Tobacco mosaic virus*
ToRSV	*Tomato ringspot virus*
TRSV	*Tobacco ringspot virus*
TRV	*Tobacco rattle virus*
TuMV	*Turnip mosaic virus*
TuYV	*Turnip yellows virus*
TYMV	*Turnip yellow mosaic virus*
WYMV	*Wheat yellow mosaic virus*

Introduction

1. General introduction on phytoviruses

Ever since humans started cultivating crops during the Neolithic period about 10 000 years ago, cultivates were confronted with adverse climatic conditions and pathogen attacks, such as viruses, fungi or bacteria. André Lwoff defined the notion of virus in 1953 using three criteria. First, a virus is a nucleoprotein entity that contains one type of nucleic acid, either DNA or RNA, which constitutes the viral genome. Second, a virus cannot undergo binary divisions and can only multiply in a cell by auto-assembly of its components. Third, a virus depends on the energy and the cellular machinery of the host cell. Viruses are obligatory intracellular parasites and they are the most abundant living entities on the planet. In this chapter, I will discuss the relevance to study phytoviruses, to better understand their replication cycle but also regarding potential agronomic applications.

1. History and agro-economic impact of plant viruses

The first plant virus that has been characterized is the tobacco mosaic virus (TMV) at the end of the XIXth century. In the 1880s, Adolf Mayer first observed cases of mosaic on tobacco leaves and concluded that it might be due to small type of bacterium (Mayer, 1968). A few years later, in 1892, Dmitri Ivanovsky further investigated this and is the first to state that this disease is not caused by a bacterium or a fungus (Ivanowski, 1968). His observations have been independently repeated by Martinus Beijerinck that has shown that the infectious agent does not replicate on a nutritive medium, instead it needs the tobacco cells to multiply. He has defined this infectious agent as the *Tobacco mosaic virus* (TMV) and used the notion of “virus” for the first time (Beijerinck, 1968). The TMV is then extensively studied and the modern virology is born.

Following this discovery, more and more viruses have been discovered and require to be classified. In 1971, David Baltimore has proposed a viral taxonomy that is still in use today. Baltimore has divided viruses into seven classes, based on the type of nucleic acid that they contain and the mode of replication of this nucleic acid. Classes I and II regroup double- and single-stranded DNA viruses, respectively. Classes III, IV and V contain RNA viruses with a double-stranded RNA (class III), a positive sense single-stranded RNA (class IV) and a negative sense single-stranded RNA (class V). Finally, class VI and VII include single-stranded RNA viruses with a DNA intermediate in their cycle for class VI, also called retroviruses, and double-stranded DNA viruses with an RNA intermediate in their cycle for class VII, alternatively named pararetroviruses (Baltimore, 1971). Viruses with a single-stranded RNA genome of positive polarity represent the most ancient group of viruses and there are more than 400 species. In plants, more than 80% of the viruses have an RNA genome (ICTV, 2022).

Plant viruses are responsible for significant damages worldwide, estimated at tens of billions U.S. dollars annually (Sastry and Zitter, 2014) and they represent half of the emerging crop diseases, which is similar to the part of human diseases caused by viruses (Anderson et al., 2004). Climate change

contributes to the emergence of viral crop diseases because it changes the distribution and abundance of arthropods vectors viruses (Anderson et al., 2004), which are the main viral vector in nature (Hogenhout et al., 2008).

2. Phytovirus transmission and hosts

Hemipteran insects like aphids transmit 55% of the vectored plant viruses (Hogenhout et al., 2008). Those insects feed on the plant using their piercing-sucking mouthparts composed of a needle-like stylet bundle (Cranston and Gullan, 2009). They can be specialized as phloem, xylem or mesophyll feeders, but many phytoviruses are transmitted by insects that are phloem-specialized (Chapman, 1998). For example, viral species belonging to the *Potyvirus* genus are exclusively transmitted by aphids (Hogenhout et al., 2008). In addition, hemipteran-transmitted viruses can be divided into two subclasses, according to the timing during which the infection is possible. Non-circulative viruses are retained in the stylet of the vector and the infection is possible only for a short time interval because the insect loses the virus within a few minutes. By contrast, circulative viruses are acquired for a longer time period and can circulate through the food canal, reaching the digestive tube of the insect. Propagative circulating viruses can replicate in the vector whereas non-propagative circulating viruses cannot. Of note, most hemiptera-transmitted viruses are non-circulative.

Nematodes are the second most common vector among viruses for which the transmission vector is identified, like the dagger nematodes from the *Xiphinema* genus. They are present in soil and while they feed on the roots, the virus is transmitted via their dagger to the phloem (Hogenhout et al., 2008). Nematode-transmitted viruses are non-circulative and stored in the oesophagus of the nematode (Fuchs et al., 2017).

A host plant is defined as a plant in which the virus can replicate. The replication can be limited to a few cells, in the inoculated leaves, or the virus can spread throughout the whole plant in “systemic” infection. There are two types of hosts: natural hosts that are naturally infected by a virus in a natural environment, and experimental hosts in which the virus can replicate when it is inoculated under laboratory conditions. Some viruses can have a relatively narrow host range in nature, but they can replicate in high number of hosts under laboratory conditions (Dawson and Hilf, 1992). The case of the TMV is well studied regarding the host range. It has a narrow natural host range but it has been shown to replicate in 199 out of 310 plant species tested under laboratory conditions (Holmes, 1946). Among these plants, some were re-classified as “nonhosts” because TMV did not propagate in the rest of the plant (Cheo, 1970, 1971; Sulzinski and Zaitlin, 1982). Interestingly, some viral species are extremely specialized, in one or two species, like the *Grapevine fanleaf virus* (GFLV) which only infects grapevine in nature. Other viral species have a much broader host range, like the *Turnip mosaic virus* (TuMV), which can infect more than 300 species (Walsh and Jenner, 2002).

2. Plant virus genomes: organization and key features

Viral genomes can be composed by various types of nucleic acids. Here, I will mainly focus on positive single-stranded RNA (ss(+) RNA) genomes of plant viruses, as they represent most of the plant

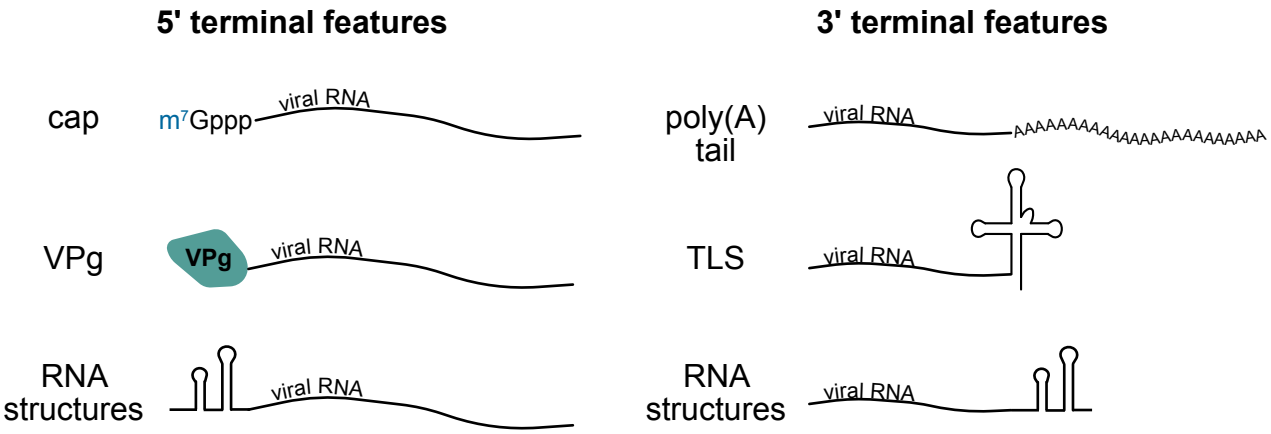


Figure 1: Recap of 5' and 3' features on viral RNAs. Viral RNA genomes can have a cap structure (cap 0 type) or a genome-linked viral protein (VPg) or RNA structures at their 5' extremity. The 3' extremity of viral RNA can be polyadenylated or fold into RNA structures that can, in some cases, mimic cellular tRNAs (tRNA-like structures, TLS).

virus genomes and our work focuses on this class of viruses. First, I will present general information on viral RNA genomes. To preserve their integrity, viral RNAs harbor features at the 5' and 3' termini that can mimic eukaryotic mRNA terminal features. I will then describe 5' and 3' features found on viral RNAs in plants, and finally, I will present cases of non-templated nucleotide addition to preserve the viral genome integrity.

1. The RNA genome: compact and efficient

RNA genomes of phytoviruses generally are small (4-12 kb) as compared to viruses infecting mammals (SARS-CoV-2 genome 30 kb), with the exception of *Closteroviruses* that can have larger genomes up to 19 kb (Agranovsky, 2021).

The single-stranded RNA genome of positive polarity can directly be recognized by the translation machinery of the plant cell. The RNA genome is polycistronic and usually does not encode for more than ten viral proteins. The "basic viral genome kit" codes for the RNA-dependent RNA polymerase (RdRp) and the structural proteins of the capsid. The viral genome sometimes also encodes helicases and proteins that help with the cell-to-cell movement or that inhibit the silencing reaction of the host, like it has been shown for the P19 protein of tombusviruses (Lakatos et al., 2004). The genome of plant RNA viruses is relatively compact and ribosomal frameshifting is often used to translate all the proteins, but the RNA can also be translated into a polyprotein that is then cleaved into several viral proteins by the viral proteinase.

Viral RNAs can be capped at their 5' extremity and polyadenylated at their 3' terminus. Other terminal features are reported for viral RNAs such as 5' genome-linked viral protein (VPg) or 3' structures, for example tRNA-like structures (TLS) or hairpin structures. 5' and 3' terminal RNA features found on viral RNAs are presented in Fig. 1. Among the 87 genera of plant-infecting viruses that have a ss(+) RNA genome, only 12 genera possess both a 5' cap and a 3' poly(A) tail. Capped viral genomes of 11 genera possess a TLS instead of a poly(A) tail. 25 genera have a 5' VPg and sometimes a poly(A) tail as well, or strong structures instead of the poly(A) tail. Of note, in 36 genera, the viral genome does not have either of the 5' cap or the 3' poly(A) tail (ICTV, 2022).

2. 5' terminal RNA features of ss(+) RNA viruses

There are three main 5' features on viral RNA genomes: a cap structure, similar to that of eukaryotic mRNAs, a genome-linked protein (VPg), bound to the first nucleotide of the viral genome, or RNA structures within the 5'UTR with no chemical structure nor protein bound to the first nucleotide. In this part, I will first introduce the eukaryotic cap structure and function, then the viral cap synthesis, the VPg characteristics, and finally examples of 5' RNA structures found in viral genomes.

a. The cap structure

mRNA cap synthesis and structure

In eukaryotes, RNAs transcribed by RNA polymerase II, including mRNAs, are modified at their 5' extremity by the addition of a structure commonly called "cap". This modification occurs in the nucleus as soon as the first 25-30 nucleotides of the newly synthesized mRNA are synthesized (Moteki and

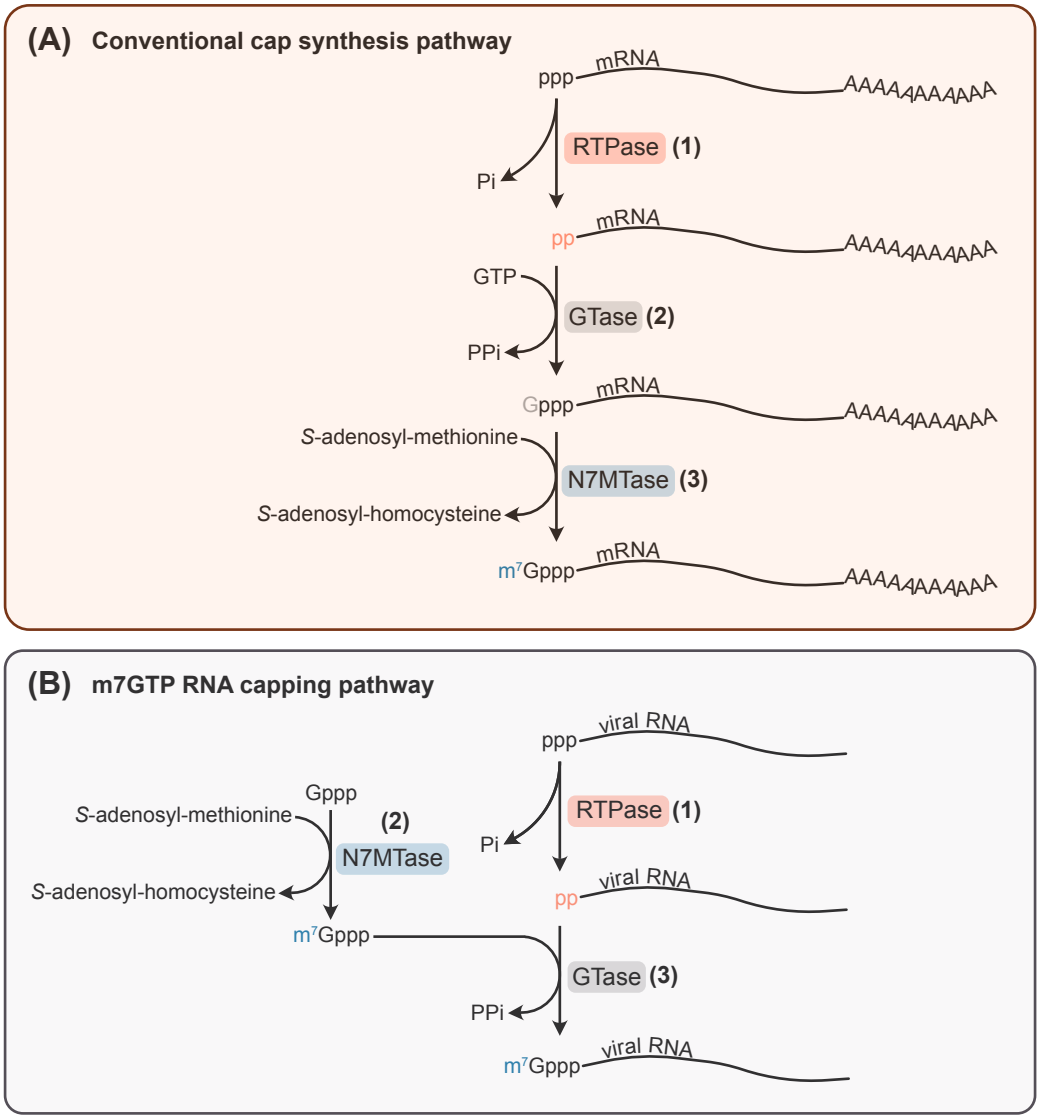


Figure 2: Cap structure synthesis pathways. (A) Conventional pathway for cap 0 synthesis on mRNAs. (1) RTPase converts the 5' triphosphate of the mRNA into 5' diphosphate. (2) GTP is used by the GTase to link a GMP molecule at the RNA 5' extremity by a 5'-5' bound. (3) The N7 amine position of this G is methylated by the MTase. **(B)** Unconventional pathway for cap synthesis on viral RNAs: m^7GTP RNA capping mechanism (mostly studied for alphaviruses but reported for some plant viruses: BaMV, BMV, TMV). (1) RTPase converts the 5' triphosphate of the mRNA into 5' diphosphate. (2) The N7 amine position of this guanine is methylated by the MTase. (3) The methylated GMP is linked at the RNA 5' extremity by a 5'-5' bound. Adapted from Decroly et *al.*, 2011

Price, 2002; Shatkin and Manley, 2000). This cap is composed of a methylated guanosine covalently linked to the first nucleotide of mRNA via a 5'-5' triphosphate linkage. The capping reaction takes place in three steps involving three enzymatic activities (Fig. 2A). The RNA triphosphatase converts the 5' triphosphate of the nascent mRNA into a 5' diphosphate. Then, a guanosine triphosphate (GTP) is used by the RNA guanylyltransferase (GTase) to add a guanosine monophosphate to the 5' diphosphate of the mRNA. The final step involves the N7-methyltransferase (MTase) that adds a methyl group to the N7 amine position of the guanine (Ghosh and Lima, 2010; Ramanathan et al., 2016). This 7-methylguanosine cap (^{m7}GpppRNA) is called cap 0.

On small nucleolar RNAs (snoRNAs) in mammalian cells, a hypermethylated version of cap 0 was found (Girard et al., 2008). Other types of caps have been detected in metazoan and eukaryotic viruses. The 2'-O position of the first transcribed nucleotide can be methylated (^{m7}GpppN$_m$pRNA, cap 1) and sometimes of the second nucleotide as well (^{m7}GpppN$_m$pN$_m$pRNA, cap 2) (Furuichi and Miura, 1975; Wei et al., 1975; Werner et al., 2011). In vertebrates, the first nucleotide of a cap 1-capped RNA can be methylated at its N6, if it is an adenosine (Mauer et al., 2017).

In the nucleus, the 5' cap structure is required during pre-mRNA processing and nuclear export. It is involved in transcription termination in human cells (Andersen et al., 2013) and is needed for efficient pre-mRNA splicing in yeast, xenopus and mammals (Fresco and Buratowski, 1996; Inoue et al., 1989; Konarska et al., 1984). In mammalian cells, the cap binding complex (CBC) binds to the cap and interacts with the U4/U5/U6 snRNP, thus initiating pre-mRNA splicing (Pabis et al., 2013). In addition, nuclear export is facilitated by the 5' cap through the CBC interaction with nuclear export factors such as ALY (Nojima et al., 2007). Once the mRNA is exported into the cytosol, the 5' cap is involved in translation. The CBC recruits the translation initiation factors eIF4E, eIF4G and eIF4A in yeast and animals (Borden and Volpon, 2020; Chiu et al., 2004; Choe et al., 2012, 2014; Fortes et al., 2000; Halstead et al., 2015).

Another important role of the 5' cap is to distinguish between self and non-self RNA. In mammalian cells, it has been shown that the 5' cap 1 structure is a signature of self RNA by abolishing the recognition of the RNA by RIG-I and LDA5, factors that trigger the type I interferon pathway (Daffis et al., 2010; Züst et al., 2011).

A 5' terminal cap structure is not limited to cap 0-2 and its derivatives. Recently, a so-called non-canonical cap was reported on eukaryotic mRNAs. NAD+-capped mRNAs in eukaryotes have first been observed in yeast, mostly on mRNA encoding translation machinery factors and mitochondrial functions (Walters et al., 2017). The same year, Jiao and colleagues have detected NAD+-capped mRNAs in mammalian cells. The NAD+-capped RNAs are enriched in small nuclear RNA (snRNAs) and small nucleolar RNAs (snoRNAs) (Jiao et al., 2017). As in Prokaryotes, the NAD+ cap is added co-transcriptionally (Walters et al., 2017). NAD+-capped RNAs have also been reported in Arabidopsis transcriptome and on the contrary to mammals, these types of caps are enriched in the polysomal fraction. Most of them are protein-encoding RNAs from the nuclear and the mitochondrial genomes (Wang et al., 2019; Zhang et al., 2019b). These RNAs have been found to be associated with active polysomes and the authors have observed a higher translation efficiency on the NAD+-capped mRNAs

(Wang et al., 2019). The roles of this non canonical cap in eukaryotic species remain to be further investigated.

Cap structure synthesis for viral RNA

Viruses belonging to the *Hepelivirales* and *Martellivirales* (except the *Endornoaviridae* family) orders, the *Alphaflexviridae*, some *Betaflexviridae* like the *Quinvirinae* subfamily, the *Vitivirus* genus and the *Tymoviridae* family (Tymovirales order), harbor a cap structure at the 5' extremity of their RNA. Among ss(+) RNA viruses in eukaryotes, only the members of the *Retroviridae* infecting vertebrates use the conventional cap synthesis pathway (Fig. 2A). Other viruses for which the replication cycle is exclusively cytoplasmic do not have access to the capping machinery within the nucleus and different mechanisms of cap synthesis have been described for them.

The final structure is the same as for mRNA, as the selection pressure is stronger on the cap structure itself than on the synthesis pathway. There are three unconventional ways to generate a cap structure on a viral RNA: the m^7GTP RNA capping, the GDP RNA capping and the RNA cap snatching pathway. I will only go into the details of the first mechanism, because the two others are mostly found in viruses with a negative sense RNA genome (Picard-Jean et al., 2013). The m^7GTP RNA capping pathway, only used by ss(+) RNA viruses, is extensively studied in alphaviruses and leads to the formation of a cap 0 structure, ^{m7}GpppRNA (Fig. 2B). The 5' triphosphate of the RNA is first hydrolyzed into a diphosphate molecule by the action of the viral RNA 5'-triphosphatase (RTPase). The viral N7-methyltransferase (N7MTase) uses S-adenosylmethionine (SAM) as methyl donor to methylate a guanosine triphosphate (GTP) molecule that is then transferred onto the 5' diphosphate by the guanylyltransferase (GTase). By contrast to conventional capping, the methylation occurs on a free GTP molecule (Decroly et al., 2011). This mechanism is conserved in plant viruses, like the tobacco mosaic virus (TMV, *Tobamovirus*) (Merits et al., 1999). In addition, similar activities were described for the bamboo mosaic virus (BaMV, *Potexvirus* (Li et al., 2001)) and for the brome mosaic virus (BMV, *Bromovirus*, (Magden et al., 2001)), suggesting a conservation of this capping mechanism through the evolution of RNA viruses. The GDP RNA capping pathway results in the formation of a cap 1 structure, ^{m7}GpppN$_m$pRNA and takes place in four steps. Finally, the RNA cap snatching is an elegant mechanism through which the cap of a host mRNA is cleaved and attached to viral mRNAs (Decroly et al., 2011).

b. The genome-linked protein (VPg)

As alternative to the 5' cap structure, viruses sometimes have a viral protein linked to the 5' extremity of their RNA, the genome-linked protein (VPg). Members of the *Potyviridae*, *Secoviridae*, *Solemoviridae* and *Tombusviridae* families harbor a VPg on their RNA. The VPg is a relatively small protein (2-24 kDa) that is covalently linked to the first nucleotide of the RNA by a phosphodiester bond between the ß-hydroxyl group of a serine or tyrosine residue and the 5' terminal nucleotide of the RNA.

This protein has various sizes, from small (2-4 kDa) for *Secoviridae*, medium size for *Sobemoviridae* and *Tombusviridae* (9-13 kDa) to large size (20-22 kDa) for *Potyviridae*. The common features of VPg proteins are a bipartite nuclear localization signal (NLS) and nucleotide triphosphate binding motifs (Jiang and Laliberté, 2011). Indeed, when expressed alone, VPg precursors localize exclusively in the nucleolus (Beauchemin et al., 2007; Rajamäki and Valkonen, 2009). Upon viral

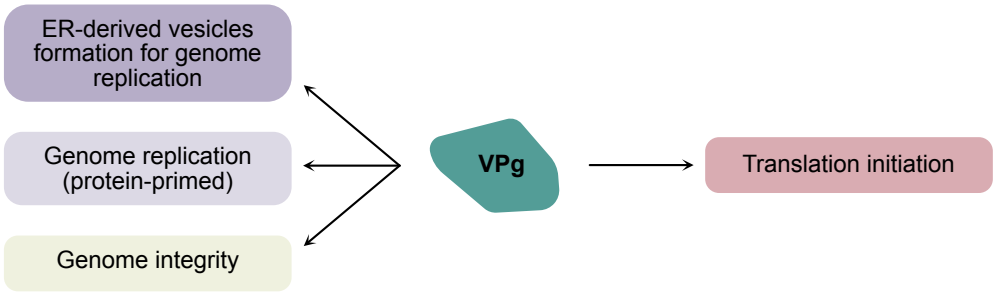


Figure 3: Functions described for the VPg. VPg on viral genomes have been reported to function in genome replication, RNA integrity, formation of ER-derived vesicles for genome replication, and translation initiation.

infection, the VPg is however only found in the cytoplasm and the viral factories (Beauchemin et al., 2007; Cotton et al., 2009). Intrinsically disordered regions have been predicted by *in silico* analyses of VPgs from 14 viral species, including species from the *Sobemovirus* and the *Potyvirus* genera (Hébrard et al., 2009). *In vitro* analyses have confirmed that VPg of potyviruses and sobemoviruses are intrinsically disordered proteins (Grzela et al., 2008; Hébrard et al., 2009; Rantalainen et al., 2008, 2011; Satheshkumar et al., 2005). The potato virus A VPg adopts a liquefied globular-like structure, partially disordered, with a hydrophobic core and a positively charged contact surface (Rantalainen et al., 2008, 2011).

The VPg is produced from a precursor polyprotein by viral proteinase cleavage. Its protein level in infected cells is mostly very low or undetectable, but VPg precursors are observed for the cocksfoot mottle virus (*Sobemovirus*, (Makinen et al., 2000)) and for the tomato ringspot virus (*Nepovirus*, (Chisholm et al., 2007)) and TuMV (*Potyvirus*, (Beauchemin et al., 2007)) at the periphery of ER membrane-derived vesicles. The VPg is proposed to be involved in the formation the ER-derived vesicles in which the viral replication takes place. Indeed, the potato virus A VPg interacts with anionic phospholipid and this interaction leads to the extension of the vesicle (Rantalainen et al., 2009).

Several other functions related to viral metabolism have been reported (Fig. 3). The disordered nature of VPg could grant the capacity to interact with many different protein partners, as it is supposed for intrinsically disordered proteins (Haynes et al., 2006). Indeed, many studies of VPg over the years tend to prove this hypothesis. For example, the VPg precursor VPg-Pro of the tobacco etch virus (*Potyvirus*) interacts with the RdRp (Li et al., 1997). The VPg of the tobacco vein mottling virus (*Potyvirus*) was reported to interact with two viral proteins, the viral protease as well as the RdRp, stimulating its activity even for a polymerase mutant with its catalytic activity attenuated (Fellers et al., 1998). In the case of the sesbania mosaic virus (*Sobemovirus*), the VPg interacts with the viral protein P10 and with the movement protein, which recognizes the viral RNA via this interaction (Chowdhury and Savithri, 2011).

VPg and its precursors also associate with host factors, in particular the translation machinery. Many articles report the interaction of the VPg with the C-terminal region of the eukaryotic initiation factor 4E (eIF4E) and its isoforms, that usually recognize the cap structure. The interaction has been reported both *in vitro* and *in planta*, for several potyviruses, including TuMV (Beauchemin et al., 2007; Duprat et al., 2002; Gallois et al., 2010; Lellis et al., 2002; Léonard et al., 2004; Okade et al., 2009; Sato et al., 2005; Schaad et al., 2000; Tavert-Roudet et al., 2017; Wittmann et al., 1997) and the tomato ringspot virus (ToRSV, *Nepovirus* (Léonard et al., 2002)). Moreover, plants defective for eIF4E are resistant to potyvirus infection (Ashby et al., 2011; Duprat et al., 2002; Gallois et al., 2010; Lellis et al., 2002; Sato et al., 2005; Yeam et al., 2007), meaning that eIF4E isoforms are required for viral multiplication.

Other translation factors interact with the VPg, like the eukaryotic initiation factor 4G (eIF4G) and its isoform iso-eIF4G. VPg of potyviruses and sobemoviruses can form a complex with eIF4E or eIF4G (Hébrard et al., 2010; Michon et al., 2006; Plante et al., 2004), thereby reducing the affinity of the eIF4E-eIF4G complex for the cap (Michon et al., 2006; Plante et al., 2004). Like eIF4E, plants deleted from eIF4G are resistant to the rice yellow mottle virus (*Sobemovirus*) in rice (Hébrard et al., 2010). Although

an interaction between the polerovirus VPg and initiation factors has not been shown, one study reports that eIFs are required for the infection of three poleroviruses (Reinbold et al., 2013).

The eukaryotic elongation factor 1A (eEF1A) interacts with both the VPg precursor VPg-Pro and the viral RdRp of TuMV in Arabidopsis (Thivierge et al., 2008). Two DEAD-box RNA helicases have been described to interact with a potyvirus VPg in *Prunus persica* and Arabidopsis, where it colocalizes with viral factories and also is an essential factor for infection (Huang et al., 2010).

Apart from factors that interact with the 5' terminus of RNA, the 6K-VPg-Pro and the VPg-Pro precursors directly interact with poly(A) binding proteins (PABP) *in planta* (Léonard et al., 2004). It has later been demonstrated that the VPg-PABP interaction leads to the relocalization of PABP in viral factories (Beauchemin et al., 2007). A recent study on TuMV describes that PABP stimulates the binding affinity of the VPg to eIF4F, subsequently activating *in vitro* translation (Khan and Goss, 2019). PABP are important, but not essential, for the replication of TuMV: in plants deleted from PABP2/4/8, three PABP isoforms most highly and broadly expressed, the virus can multiply but the viral RNA accumulates less. Upon infection, the mRNA and protein levels of Arabidopsis PABP2, PABP4 and PABP8 increases. The authors show that, *in vitro*, the viral RdRp and the VPg interact directly with the PABPs, with PABP2 being their favorite interactant (Dufresne et al., 2008a).

Like the cap structure, the VPg is involved in the protection of the 5' extremity and the translation of the viral RNA, but also viral RNA replication. It can be used as primer to initiate RNA synthesis. One of the most known examples is the replication mechanism of the poliovirus (*Picornaviridae*), a virus infecting mammals. Upon the combined action of RNA structures, a viral RNA binding protein and the viral RdRp, the VPg is uridylylated and the resulting VPg-pUpU is used by the RdRp as protein primer (Paul and Wimmer, 2015). In plant ss(+) RNA viruses, protein-primer replication via the VPg has been reported for two viruses from the *Potyviridae* family, the potato virus A and the peppervein banding virus (Anindya et al., 2005; Puustinen and Mäkinen, 2004). In both cases, the VPg is uridylylated like for the poliovirus: this suggests similarities in the replication mechanism between the *Potyviridae* and *Picornaviridae* families.

c. 5' RNA structures

Some ss(+) RNA viruses do not harbor a chemical structure nor a protein at the 5' extremity of their RNA, which 5' extremity is a 5' triphosphate. In this case, the 5'UTR folds into secondary structures and contains elements that helps the recruitment of ribosomal subunits to start the translation in a cap-independent manner (Geng et al., 2021; Kneller et al., 2006). For example, the *Turnip crinkle virus* (TCV, *Betacarmovirus*) 5'UTR contains translation enhancer sequences that promotes the recruitment of ribosomal subunits (Qu and Morris, 2000).

3. 3' terminal RNA features of ss(+) RNA viruses

At their 3' extremities, viral RNA genomes can be polyadenylated, like eukaryotic mRNAs, or the 3'UTR can fold into RNA structures. Here, I will first provide information on the mRNA poly(A) tail synthesis and function, then on viral RNA polyadenylation and finally, I will present the structures found on viral RNAs: tRNA-like structures (TLS) and other types of RNA structures, as well as their identified functions in viral metabolism.

a. The poly(A) tail

Poly(A) tail synthesis for mRNAs

Messenger RNAs are decorated at their 3' termini by a poly(A) tail. In fact, all transcripts emerging from RNA polymerase II transcription are polyadenylated. The addition of adenosines is ensured by poly(A) polymerases (PAP), containing a nucleotidyl transferase (NTr) and a PAP-associated domains (De Almeida et al., 2018a; Warkocki et al., 2018a). These enzymes belong to the class I of TNTases, together with the archaeal tRNA CCA-adding enzymes, the 2'-5'-oligo(A) synthetases and terminal nucleotidyl transferases adding adenosines, uridines, cytosines or guanosines (Jae Eun and Wickens, 2007; Martin and Keller, 2007).

In Arabidopsis genome, four genes code for PAPs. PAPS1, PAPS2 and PAPS4 are localized in the nucleus (Meeks et al., 2009) and constitutively expressed (Hunt et al., 2008). They are responsible for the co-transcriptional addition of the poly(A) tail on RNA polymerase II transcripts. Mutations in these genes induce growth phenotypes such as reduced leaf size, male gametophyte defect (Vi et al., 2013) or early flowering (Czesnick and Lenhard, 2016). Several studies on Arabidopsis PAPs have shown that these enzyme target different populations of transcripts, for example coding for ribosomal proteins, auxin signaling, flowering regulators and oxidative stress response, suggesting Arabidopsis PAPs are functionally specialized (Czesnick and Lenhard, 2016; Kappel et al., 2015; Trost et al., 2014; Vi et al., 2013). On the contrary to the other PAPs of Arabidopsis, PAPS3 lacks the C-terminal part and its expression pattern is restricted to pollen (Hunt et al., 2008). PAPS3 is localized in the cytosol and its function is not determined yet (Meeks et al., 2009).

First reported in the early 1970s, the poly(A) tail has been proposed to be involved in mRNA transport or translation (Passmore and Coller, 2021). The synthesis of the poly(A) tail is mostly studied in metazoans and requires many factors. These factors include the cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulation factor (CSTF), the cleavage factor I (CFI), the cleavage factor II (CFII), the poly(A) polymerase, the nuclear poly(A) binding proteins (PABPN) and the scaffold protein symplekin. The mRNAs contain a polyadenylation signal with a canonical sequence AAUAAA downstream of a GU- or U-rich sequence that promotes the recruitment of these factors (Tudek et al., 2018) The RNA is cleaved during the transcription 10-30 nt downstream of the polyadenylation signal and the addition of adenosines is catalyzed by a canonical poly(A) polymerase (cPAP). Once 11-14 adenosines have been added, PABPN binds to the poly(A) tail (Meyer et al., 2002). After the binding of a PABPN, the cPAP switches from distributive to processive mode and quickly synthesize to full length poly(A) tail. In the nucleus of human cells, poly(A) tail length is 200 adenosines and can reach up to 400 As (Nicholson-Shaw et al., 2022). In the cytosol, the median length of the poly(A) tail is shorter, within the range of 50-100 adenosines in *H. sapiens, M. musculus, D. melanogaster, C. elegans and A. thaliana*, and about 30 adenosines in yeast (Chang et al., 2014; Eisen et al., 2020; Lima et al., 2017; Subtelny et al., 2014; Workman et al., 2019). The poly(A) tail is coated by cytoplasmic PABP (PABPC) that repetitively binds to the poly(A) tail with a footprint of 20-30 As (Lima et al., 2017; Yi et al., 2018).

In plants, the elements required to recruit the polyadenylation complex differ from their counterparts in yeast and mammals. The near-upstream element (NUE) is located between 10-40 nt upstream of the

polyadenylation site and is characterized by an A-rich motif of 6-10 nt long. The far-upstream element (FUE) is located further upstream of the polyadenylation sequence, about 100 nt, and the base composition resembles those of the downstream sequences found in mammals. Together with the polyadenylation site itself and its adjacent U-rich motif, NUE and FUE form the signal necessary for 3' polyadenylation of plant mRNAs (Hunt et al., 2008). Orthologs of metazoan polyadenylation factors have been identified in plant species (Hunt et al., 2012). CPSF and CSTF mutants in plants exhibit development alterations, like a delayed flowering time (Liu et al., 2010, 2014) and alterations in the choice of the polyadenylation site (Liu et al., 2014; Thomas et al., 2012; Zeng et al., 2019). On the contrary to animals, null mutants of these factors are viable. An explanation to this unexpected finding could be that plant genomes contain other genes that may substitute them and/or act redundantly (Hunt, 2020).

Poly(A) binding proteins

Poly(A) tails are coated by poly(A) binding proteins (PABP), either in the nucleoplasm (PABPN) or in the cytosol (PABPC). These proteins bind to oligo-adenosines with high affinity (Görlach et al., 1994) via the RNA-recognition motif (RRM) domains as it has been reported for human PABPN1 (Kerwitz et al., 2003). PABPN1 has one RRM domain and a C-terminal arginine-rich domain (Mangus et al., 2003). One of the functions of nuclear PABP binding may be to protect RNA 3' termini from degradation by 3'-5' exoribonucleases. In human cells, transcripts associated with PABPN1 tend to have longer tails than those bound by PABPC, independently of their cellular localization, so the association with PABPN1 may be a requirement to maintain long poly(A) tails (Nicholson-Shaw et al., 2022). Human PABPN1 association to the poly(A) tail allows to improve the affinity of the cPAP for RNA (Kerwitz et al., 2003) to promote polyadenylation.

By contrast, an interaction between the yeast PABPN Nab2 and the enzymatically active subunit of the exosome Rrp6 has been reported *in vitro* and *in vivo* (Schmid et al., 2012). This has been further demonstrated in human cells with the discovery of the Poly(A) tail eXosome Targeting (PAXT) connection involving the interaction between PABPN1 and the helicase MTR4, via ZCF3H1, a zinc-finger protein (Meola et al., 2016). These data provide evidence for the role of PABPN in the regulation of the abundance of pre-mRNA in the nucleus.

There is not much information on how the transition between nuclear and cytoplasmic PABP coating occurs. One study proposes that it is triggered by a first round of translation (Sato and Maquat, 2009), but the nuclear export could also play a role in the switch. PABPN-coated poly(A) tails do not present the repeating PABP footprint (Lima et al., 2017; Yi et al., 2018) and they are longer as compared to PABPC-covered poly(A) tails (Nicholson-Shaw et al., 2022). Moreover, transcripts associated with PABPC have a higher translation efficiency than those detected in the PABPN1 immunoprecipitation (Nicholson-Shaw et al., 2022). These data suggest that poly(A) tails bound to PABPN may be less accessible to deadenylases and less translated, thereby probably not present in the cytoplasm.

PABPC are evolutionary conserved in eukaryotes. Multiple PABP-encoding genes are present in metazoan and plant genomes (Mangus et al., 2003). In human cells, there are four PABPs, one is located within the nucleus and the other three within the cytoplasm. In plants, only PABPN1 of *Citrus sinensis* has been reported to bind poly(A) tails (Domingues et al., 2015) and nuclear PABP remain

uncharacterized. In Arabidopsis, eight genes code for cytoplasmic PABPs (Goss and Kleiman, 2013). PAB2, PAB4 and PAB8 are the most widely and highly expressed, while PAB6 and PAB7 are expressed at a lower level. PAB3 and PAB5 expression is restricted to reproductive tissues. PAB1 has a weak tissue-specific expression (Belostotsky, 2003). These patterns of expression seem to be conserved in *Oryza sativa* (Siddiqui et al., 2007).

A PABP is composed of four RNA-recognition motifs (RRMs) and the C-terminal domain contains a PABC domain involved in protein-protein interaction (Sachs et al., 1986). The RRMs are responsible for the binding to the poly(A) tail and a minimum of 8 adenosines is required for the binding of one RRM (Webster et al., 2018). In plants however, PAB1 and PAB6 do not have four RRM domains, so it has been proposed that the duplicated copies of PABP genes in plants may have evolved towards specialized functions in mRNA metabolism (Belostotsky, 2003; Siddiqui et al., 2007).

The binding of PABPC prevents RNA degradation from its 3' extremities *in vitro* (Bernstein et al., 1989). *In vivo*, the phasing pattern observed for poly(A) tails suggests that uncoated poly(A) tails are more prone to be deadenylated (Lima et al., 2017; Yi et al., 2018). In addition, PABPC interacts with translation factors eIF4G and eRF3 (Hoshino et al., 1999; Tarun et al., 1997), thereby stimulating translation.

However, PABPC proteins have a dual role because several studies report that they are involved in the positive regulation of deadenylation. Longer poly(A) tails have indeed been observed upon the depletion of Pab1 in yeast (Caponigro and Parker, 1995). In addition, PABPC have been demonstrated to directly interact with the deadenylases PAN2/3 and the CCR4/NOT complex in yeast and human cell lines (Funakoshi et al., 2007; Uchida et al., 2004; Webster et al., 2018). In human cells and yeast, the CCR4 deadenylase can release PABPC from the poly(A) tail and further deadenylate the RNA (Webster et al., 2018; Yi et al., 2018). In addition, PABPC binding to the poly(A) tail hinders deadenylation by CAF1 but not by CCR4 (Webster et al., 2018; Yi et al., 2018). Interestingly, it has been shown in yeast that PABPC can bind to 3'UTR, in particular on AU-rich sequences (Webster et al., 2018).

Altogether, the PABP are key players of the mRNA metabolism, involved at every step of a mRNA life cycle.

Viral RNA polyadenylation

Like mRNAs, viral RNAs can have a 3' poly(A) tail. So far, the number of studies on plant viruses is limited, so I will expose some examples from metazoan viruses in addition to what is published on phytoviruses. Several key questions lie in the mechanism of poly(A) tail synthesis, especially for viruses that only are cytoplasmic: do they usurp the cellular machinery for poly(A) tail synthesis and cause its relocalization into the cytoplasm? Do they possess their own polyadenylation activity? Are they polyadenylated by an unidentified cytoplasmic poly(A) polymerase? There are three strategies proposed so far. First, the 3'UTR of the viral RNA can contain a polyadenylation signal that will, like for mRNAs, regulate the polyadenylation site and subsequent adenosines addition. This would require to hijack the polyadenylation machinery from the nucleus and to relocalize it within the cytoplasm. Alternatively, the viral polymerase could ensure the adenosines addition or a cellular poly(A) polymerase could be usurped by the virus. Finally, another usually accepted hypothesis lies in the mechanism of viral RdRp stuttering on a homopolymeric sequence.

The presence of a putative and/or confirmed polyadenylation signal in the 3'UTR of animal viral RNAs was reported for several viral species from the *Enterovirus* and *Coronavirus* genera and the Sindbis virus (*Alphavirus*) (van Ooij et al., 2006; Peng et al., 2016; Raju et al., 1999). In addition, the viral RdRp can synthetize the poly(A) tail decorating the 3' extremity of the viral RNA using a poly(U) stretch of the template strand. This strategy was reported for alphaviruses, coronaviruses and picornaviruses (Kempf and Barton, 2015; Sawicki and Gomatos, 1976; Steil et al., 2010; Wu et al., 2013b). The viral RdRp stuttering on a homopolymeric tract has been described for negative sense RNA viruses, belonging to the *Rhabdovirus* (Barr et al., 2002), *Filovirus* (Volchkova et al., 2015), *Orthomyxovirus* (Poon et al., 1999; Zheng et al., 1999) and *Paramyxovirus* genera (Hausmann et al., 1999). In plants, polyadenylation sites have been described for *Potexvirus* species and the *Rice tungros bacilliform pararetrovirus* (Chen et al., 2005; Guilford et al., 1991; Rothnie et al., 2001). The polyadenylation activity of the replicase complex from the *Bamboo mosaic potexvirus* has been demonstrated *in vitro* and it can synthetize poly(A) tails up to 200 residues (Chen et al., 2013).

Poly(A) tail lengths of viral RNAs were first determined in 1979 for a few viruses infecting animals and plants. Using 3'-end ^{32}P-labelling and ribonuclease digestions, the authors show that the poly(A) tails display extensive variations in length (Ahlquist and Kaesberg, 1979). Over the years, other studies have shown that the poly(A) tail lengths vary between the viruses as it was observed in picornaviruses (Kempf and Barton, 2015), coronaviruses (Shien et al., 2014) and the *Wheat yellow mosaic virus* belonging to the *Potyviridae* family (Geng et al., 2019). The poly(A) tail length has also been proposed to vary during viral infection. In a study with the bovine coronavirus, the authors observed that the poly(A) tail of viral RNAs increases during early infection and then gradually decreases from 12 hours post-infection. The authors proposed that the poly(A) tail length is regulated over the course of infection (Wu et al., 2013b). However, they do not say whether the poly(A) tail shortening is accompanied by a destabilization of the viral RNA, caused by a progressive deadenylation and subsequent degradation of the viral RNA. Data in metazoans tend to indicate that the viral polyadenylation could be regulated by alternative polyadenylation in the case of the glycoprotein-coding gene of the Ebola virus (Volchkova et al., 2015).

Viral RNA polyadenylation has an impact on host mRNA polyadenylation and polyadenylation-related processes, as it has been reported by a few studies in animals (Vijayakumar et al., 2022). The influenza A virus NS1 protein hinders pre-mRNA polyadenylation by binding to the cellular poly(A) binding protein 2 (PABP2) (Chen et al., 1999; de Rozières and Joseph, 2020) and inhibits the cleavage of Hsp70 pre-mRNA at the polyadenylation site (Shimizu et al., 1999). Moreover, the polyadenylation sites of cellular mRNAs can be altered upon alphavirus infection. The protein HuR, involved in the regulation of alternative polyadenylation, is relocalized from the nucleus into the cytoplasm, causing a massive dysregulation of the polyadenylation sites (Barnhart et al., 2013; Dickson et al., 2012). A more recent study also reports the alteration of alternative polyadenylation upon vesicular stomatitis virus infection (Jia et al., 2017). Of course, these alterations may reflect on the general expression pattern of immune response factor-encoding genes (Cross et al., 2019), leading to the deregulation of the antiviral response.

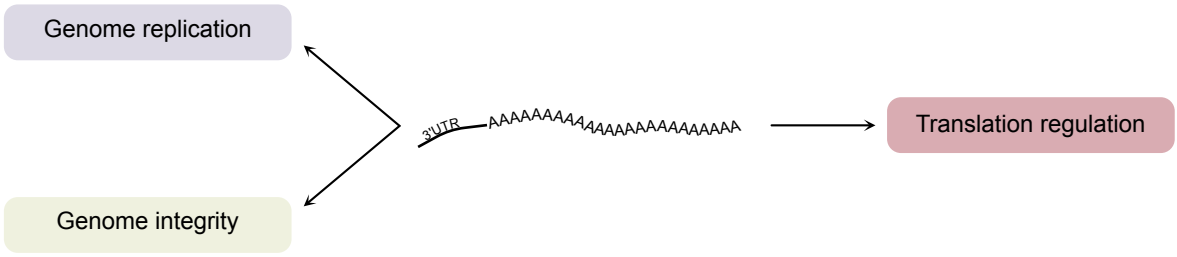


Figure 4: Functions of the poly(A) tail on viral RNAs. The poly(A) tail of viral RNAs has been described to be involved in genome replication and integrity, as well as in translation regulation.

Like for mRNAs, the poly(A) tail of viral RNAs is involved in translation regulation (Fig. 4). A recent *in vitro* study on the WYMV, which RNA is polyadenylated, shows that the 5' and 3'UTRs and the poly(A) tails of the viral RNA promote translation of a reporter gene. The addition of more adenosines residues to the 3' extremity had an opposite effect on translation, so RNAs with long tails were less efficiently translated (Geng et al., 2019). This *in vitro* experiment provides promising results on the role of poly(A) tails on viral RNAs but *in vivo* data is lacking to fully understand its function.

The poly(A) tail of mRNAs has been described to be able to interact with elements within the RNA, establishing structures that prevent 3'-5' decay. Sequences involved in this interaction have been named element for nuclear expression (ENE) and they prevent RNA decay via triple helix formation with the poly(A) tail (Conrad et al., 2006; Torabi et al., 2021a). An ENE has been described for a noncoding RNA produced by Kaposi's sarcoma-associated herpesvirus (Mitton-Fry et al., 2010). More recently, ENEs and potential poly(A) tail interaction have also been predicted in viral RNAs from the *Picornavirales* order (Torabi et al., 2021b). To date, there is no data regarding plant viruses, but the presence of similar elements is not excluded and would be worth exploring. Very recently, analyses on Sindbis virus RNA (*Alphavirus*) reports that the poly(A) is involved in the formation of a pseudoknot structure (Olsthoorn, 2022), similarly to what has been reported for the plant virus BaMV (*Potexvirus*) (Tsai et al., 1999). Even though the paper on Sindbis virus has not been peer-reviewed yet, it brings promising results regarding another role of the poly(A) tail on viral RNAs.

The poly(A) tail on viral RNA can also serve as template to initiate the minus strand synthesis during the viral genome replication (Fig. 4). Hence, for many viruses, the minus strand begins with a poly(U) stretch, complementary to a part of the poly(A) tail. For the poliovirus (*Picornaviridae*) in animals, it has even been shown *in vitro* that the minimal length of the poly(A) tail required for viral replication is between 12 and 20 residues (Silvestri et al., 2006). Moreover, the replication of the polioviral RNAs has now been thoroughly described and relies on a poly(A) at the 3' extremity of the (+) strand, a poly(U) at the 5' end of the (-) strand and the VPg (Kempf and Barton, 2015; Steil and Barton, 2009; Steil et al., 2010). The Sindbis virus genome minus strand also has a poly(U) tract that is proposed to have a similar function (Hill et al., 1997). In plants, the bamboo mosaic potexvirus RNA replication was described to be initiated from multiple sites within the poly(A) tail. The minus strand begins with a poly(U) tract as well (Cheng et al., 2002).

b. tRNA-like structures

Alternatively to the classical poly(A) tail, 3'UTRs of viral RNAs can fold into highly structured architectures that can mimic a cellular tRNA and are called tRNA-like structures (TLSs). TLSs have been reported for many genera of plant-infecting viruses: *Bromovirus, Cucumovirus, Furovirus, Hordeivirus, Pecluvirus, Pomovirus, Tobamovirus* and *Tymovirus*. They are structured as a L-shape folded around a pseudoknotted aminoacyl acceptor arm. To be categorized as TLS, it is often considered that the 3' structure must be adenylated by the cellular CCA-adding enzyme and/or have an aminoacylation capacity and/or the ability to bind with the elongation factor 1 in complex with GTP (Dreher, 2010). If the viral genome is multipartite, each genome segment has a TLS and those are identical or almost indistinguishable, as it has been shown for bromoviral and cucumber mosaic virus TLSs (Ahlquist et al., 1981). Moreover, some viral proteins are translated from subgenomic RNAs

(sgRNAs) that also possess a TLS. In plants, the TLS can be esterified either with valine or histidine or tyrosine (Dreher, 2010; Wu et al., 2021).

The valine-charged TLSs are mostly found in *Tymovirus* genomes and they have the closest structure to a cellular tRNA. A recent study identified 108 new putative TLSs that could be aminoacylated with valine in 46 different viruses, many among *Tymoviridae* but some belong to the *Virgaviridae* family and others even are insect-infecting viruses (Sherlock et al., 2021). The turnip yellow mosaic virus (TYMV) TLS valylation has been the first to be discovered (Pinck et al., 1970; Yot et al., 1970) and since then, TYMV TLS structure and function have been extensively studied. Prediction and nucleotide probing allowed to propose its structure in 1983 (Rietveld et al., 1983). *In vitro*, TYMV TLS harbors the three tRNA properties that we previously mentioned: it can be adenylated, aminoacylated and is able to bind to eEF1A.GTP (Dreher and Goodwin, 1998; Litvak et al., 1973). It is likely that TLS similar to TYMV also share these properties, although there is not much data for others so far (Dreher, 2010). Because most TYMV RNAs end with 3'CC, it has been proposed that 3' adenylation is required before its aminoacylation, like cellular tRNAs (Giegé et al., 1978). Other *Tymovirus* and some *Virgaviridae* TLS can be aminoacylated with valine but they cannot bind to the eEF1A.GTP complex as efficiently as TYMV TLS (Dreher and Goodwin, 1998; Goodwin and Dreher, 1998). Indeed, the peanut clump virus (PCV) and the indian peanut clump virus (IPCV), both from the *Pecluvirus* genus, have an insertion within the TLS as compared to the TYMV TLS. Of note, in the sequence of PCV RNA2 tested by Goodwin and Dreher, the TLS lacks a nucleotide pair required for valylation specificity, so the aminoacylation efficiency is much lower (Goodwin and Dreher, 1998), which is contradictory to a more recent aminoacylation test on this TLS (Sherlock et al., 2021).

TLSs found in *Tobamovirus* genomes are charged with histidine (Dreher, 2010). The secondary structure of TMV TLS has been proposed by Rietveld and colleagues in 1984, based on prediction and nucleotide probing (Rietveld et al., 1984). They also studied TLS of other tobamoviruses and they show that all of the TLSs studied are substrates of the cellular CCA-adding enzyme *in vitro*. Their structure differs from those of cellular tRNA so the histidylation of tobamovirus TLS relies on the histidine identity elements found within the TLS *i.e.,* the set of nucleotides responsible for the specific aminoacylation, rather than its structure (Rudinger et al., 1997). TMV TLS also meets the three properties of a tRNA since it was reported to interact with eEF1A, in an aminoacylation-dependent manner (Litvak et al., 1973). The TLS of one *Tymovirus* species, the Nemesia ring necrosis virus, has also been described to be esterified with histidine (Koenig et al., 2005), by contrast with other tymoviruses where valine is used for this reaction.

Tyrosine-charged TLSs are characteristic of bromoviruses, in particular the *Brome mosaic virus* (BMV) (Dreher, 2010). The model for BMV TLS has been established over the years (Dreher and Hall, 1988; Fechter et al., 2001; Felden et al., 1994; Rietveld et al., 1983) and like the histidine-charged TLS, the tyrosine identity elements are essential for the tyrosylation of BMV TLS (Dreher, 2010). BMV TLS can also be recognized by the CCA-adding enzyme et be adenylated (Dreher and Hall, 1988). Once aminoacylated, BMV TLS can interact with EF1A (Bastin and Hall, 1976). This is not broadly known for viral TLSs, but in the case of BMV and another bromovirus, the barley stripe mosaic virus, none of the encapsidated RNAs are aminoacylated (Loesch-Fries and Hall, 1982).

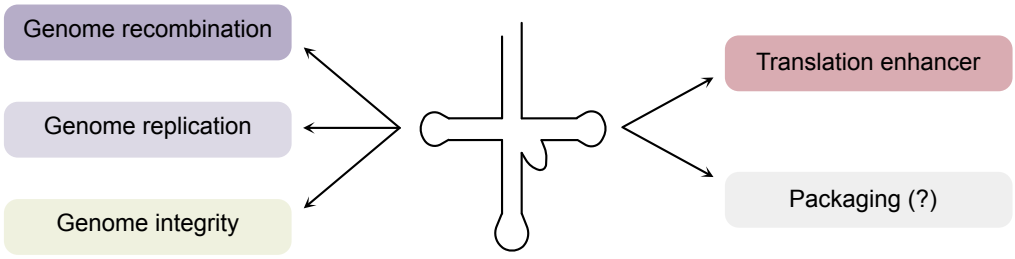


Figure 5: Functions described for plant virus TLS. tRNA-like structures in phytoviruses have been reported to function in genome replication, RNA integrity, genome recombination, translation regulation. A role in viral genome packaging has been proposed. The cloverleaf structure represents the TLS.

Of note, the tobacco rattle virus (TRV, *Tobravirus*) RNAs present a 3' terminal structure that resembles that of tymo- and tobamoviruses. Even though the RNA2 can be a substrate for the cellular CCA-adding enzyme, none of the twenty common amino acids can be charged onto this structure. There is no experimental data on RNA1 so far but the 3'UTR sequences of RNA1 and RNA2 are identical so the authors proposed that they share the same properties (van Belkum et al., 1987). So, the fact that TRV RNAs do possess a TLS is debatable.

The many studies on TLS functions tend to show that there is no standalone function for all the viral genomes that possess a TLS (Fig. 5). Indeed, the functions vary between TLS and viruses. A striking example lies in two valine-charged TLS, TYMV and PCV TLSs. The aminoacylation of TYMV TLS is essential for its infectivity (Tsai and Dreher, 1991), whereas it is not for PCV and only confers a mild fitness advantage *in planta* (Matsuda et al., 2000). The efficiency of infection related to the TLS aminoacylation can rely on several factors like a better RNA stability or replication, or a more efficient translation. All of these roles have been investigated and been reported for viral TLSs in plants. TLSs can have telomere-like functions, thereby maintaining the genome integrity and protecting against nuclease trimming. Experiments on BMV TLS with the RNA3 lacking a few 3' terminal nucleotides have shown that they are viable and even that the mature 3' ends are restored *in vivo* (Rao et al., 1989). The 3' adenylation that occurs in the host cell is required for an efficient replication of the genome for TYMV and BMV (Deiman et al., 1998; Dreher and Hall, 1988; Singh and Dreher, 1998). The minus strand synthesis initiation has first been thought to require the whole TLS, that would be used as primer. However, TYMV and BMV genome replication initiation simply requires the last three or four nucleotides, respectively (Chapman and Kao, 1999; Deiman et al., 1998; Dreher and Hall, 1988; Dreher et al., 1984; Singh and Dreher, 1998). To date, the connection between aminoacylation and replication initiation promotion is not determined. However, TLSs do regulate the access to the 3' replication initiation site as it has been shown for TYMV via the binding of eEF1A (Matsuda and Dreher, 2004).

TLSs have been proposed to act as translational enhancer, notably for BMV and TYMV even if this function remains controversial. For BMV, a decrease in the *in vitro* translation has been reported for BMV RNA lacking its TLS (Barends et al., 2004). For TYMV, *in vitro* experiments on a reporter gene have demonstrated that translation enhancement also requires aminoacylation (Matsuda and Dreher, 2004). Translation enhancement by the TLS has been proposed to result from the ribosome recruitment onto this structure and the TLS would be involved in the initiation step of viral protein synthesis. The authors show that the ribosome interacts with the 3' valylated-TLS of TYMV RNA and it results in a viral polyprotein starting with a valine residue at its N-terminal extremity. When the TLS is disrupted, the polyprotein synthesis is abolished and the addition of TYMV TLS *in trans* can restore the synthesis (Barends et al., 2003). Yet, these results have been contradicted as in another study, the authors report that translation is cap-dependent and that TLS deletion does not affect the translation efficiency *in vitro* (Matsuda and Dreher, 2007). In line with these results, the investigation of TMV (*Tobamovirus*) and NeRNV (*Tymovirus*) TLS functions in translation have shown that the amino acid on the TLS is not incorporated in the nascent protein and that the lack of TLS does not change the translation efficiency (Rudinger-Thirion et al., 2006). To my knowledge, the role of TLS in translation promotion is not elucidated in more viral TLSs.

Of note, viral TLSs have been proposed to play a role during RNA recombination of BMV genome (Bujarski and Kaesberg, 1986; Kao and Ahlquist, 1992; Ni et al., 2014). In addition, the packaging of the viral genome can depend on the TLS. *In vitro*, BMV genome encapsidation depends on the presence of the TLS, or of non-viral tRNAs that can also ensure the encapsidation in cooperation with the capsid protein as well (Choi et al., 2002). By contrast, experiments on another member of the *Bromoviridae* family, the *Cowpea chlorotic mottle virus*, show that the TLS is not needed for virion assembly. The capsid protein of this virus can even package other viral RNAs that do not have a TLS (Annamalai and Rao, 2005).

c. Other RNA structures

Some plant viral RNAs can present other structures in their 3'UTR as well, some of them resembling TLSs. For example, stem-loop structures have been described in the 3'UTR of alfamoviruses and ilarviruses (*Bromoviridae*). The proposed model looks like the TLS found in bromovirus genomes (Olsthoorn et al., 1999). Interestingly, this article shows that the 3'UTR of alfamoviruses and ilarviruses can adopt two mutually exclusive conformations. In one case, the coat protein binds with high affinity to the stem-loops, thereby potentially promoting translation. Alternatively, it folds into a pseudoknot structure, that could help initiate the minus strand synthesis. Another study on an *Ilarvirus* species has been published a few years later and describes a similar mechanism (Aparicio et al., 2003). Yet, the switch-conformation model of alfamoviruses and ilarviruses replication has been challenged a few years later (Petrillo et al., 2005). In this study, the authors also argue that the pseudoknot conformation of AIMV 3'UTR cannot be considered as a TLS because it lacks the 3' CCA terminus, it is a poor substrate of the CCA-adding enzyme and it cannot be aminoacylated (Olsthoorn et al., 1999).

In tymoviruses, dulcamara mottle virus (DuMV) RNA does not possess a 3' TLS. By contrast with all tymoviral RNAs described so far, the 3'UTR of DuMV RNA ends with UUC-3'. In addition, the 3'UTR is rather long and contains a poly(A) tract, feature that has never been previously reported in tymoviral RNA 3'UTRs (Tzanetakis et al., 2009).

One of the best characterized examples in terms of 3'UTR structure is probably the 3'UTR of the turnip crinkle virus genome (TCV, *Betacarmovirus, Tombusviridae*). TCV RNA 3'UTR is structured into hairpins that can interact and fold into a larger structure, and it ends with a six nucleotide extension ($CUGCCC_{OH}$-3') (McCormack et al., 2008; Simon, 2015). There are six hairpins: M3H, also called "unstructured" region (USR) because it is weakly stable (Carpenter et al., 1995), H4, H4a, H4b, H5 and Pr. Hairpins H4a, H4b and H5 fold into a T-shaped structure (TSS) and a bulge of H5 also interacts with the 3'-extension (Simon, 2015) (Fig. 6). The TSS shares structural similarities with a canonical tRNA and binds to ribosomes (Stupina et al., 2008). It is proposed to act as 3' cap-independent translation enhancer (3'CITE) but the mechanism remains unclear (Simon, 2015). As for TLSs, these strong structured 3' features protect against nuclease trimming and help in the replication and translation of the viral genome.

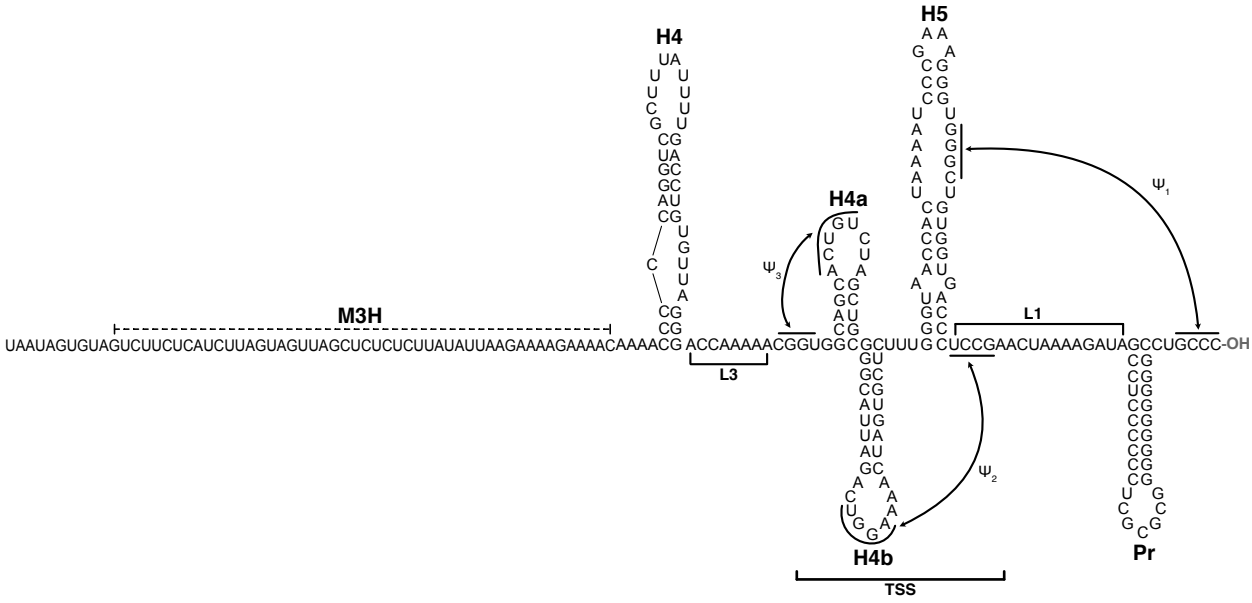


Figure 6: Secondary structure and tertiary interactions in the 3'UTR of TCV RNA. Nucleotides from nt3803 to nt4050 (last nucleotide) of TCV RNA reference sequence are shown. TCV RNA 3'UTR is composed of a weakly structured hairpin (M3H, dotted line) and five stable hairpins (H4, H4a, H4b, H5, Pr). Three pseudoknots have also been described (Ψ_1, Ψ_2, Ψ_3), indicated by the arrows. A T-shaped structure between hairpins H4a, H4b, and H5 has been described. Adapted from Simon, 2015.

4. 3' non-templated nucleotides addition and genome integrity

a. Poly(A) tail restoration

Viral RNAs present in the cytosol are targeted by nucleases that can trim the RNA extremities if they are accessible. As the viral RNA is the support of the genetic information, it is crucial that its integrity is maintained over time, after several replication cycles and/or nuclease attacks. Interestingly, cases of genome repair by addition of non-templated nucleotides have been described for viruses infecting animals and plants. *In vivo* restoration of the poly(A) tail has been reported for the poliovirus (Neufeld et al., 1994) and the hepatitis A virus (Kusov et al., 2005), the rubella virus (Chen and Frey, 1999) and the Sindbis virus (*Togaviridae*) (Hill et al., 1997; Tomar et al., 2006).

In plants, experiments on infectious transcripts lacking poly(A) tails for the beet necrotic yellow vein virus (*Benyviridae*), two *Potyviridae*, the clover yellow vein virus and the plum pox virus, and the clover mosaic virus (*Alphaflexviridae*) have reported *in vivo* poly(A) tail restoration (Guilford et al., 1991; Jupin et al., 1990; Riechmann et al., 1990; Tacahashi and Uyeda, 1999). For the clover mosaic potexvirus, the authors propose that the poly(A) tail restoration requires an AAUAAA motif, similar to the mRNA polyadenylation signal. In this study, Guilford and colleagues also show that the presence of nucleotides different from adenosines in the 3' tail abolishes the infectivity (Guilford et al., 1991). After inoculation of *in vitro* transcripts of the plum pox potyvirus under the control of a T7 promoter and lacking a poly(A), the viral progeny *in planta* has been reported to have a poly(A) tail on its RNA. In addition, the extra guanosines synthesized as first two nucleotides after an *in vitro* transcription with a T7 promoter have been lost during viral replication. These data show that the viral RNA poly(A) tail can be restored *in vivo* and that RNA extremities can be optimized during viral replication (Riechmann et al., 1990). The repair of the poly(A) tail has also been observed for another potyvirus, the clover yellow vein virus (Tacahashi and Uyeda, 1999). These results also suggest that viral RNAs can be optimized *in vivo* after several replication cycles. *In vivo* optimization of the viral RNA extremities has been described for the cowpea mosaic virus (*Comovirus*), as well as the restoration of the poly(A) tail (Eggen et al., 1989).

b. Mixed tailing on viral RNAs

After 3' extremity repair, the 3' tail can contain uridine-stretches or AU-rich sequences. First observed on the beet necrotic yellow vein virus (BNYVV, *Benyvirus*), this mechanism results in an U-rich sequence followed by the poly(A) tail (Jupin et al., 1990). It was next reported for human viruses such as the Sindbis virus (Raju et al., 1999), the coxsackie B virus (van Ooij et al., 2006) and the hepatitis C virus (van Leeuwen et al., 2006). Of note, the hepatitis C virus RNA is not described as polyadenylated.

For the Sindbis virus, Raju and colleagues proposed that a TNTase activity is encoded by the viral genome. Indeed, in some cases, the non-templated nucleotide addition depends on the viral RdRp and its active site (Poranen et al., 2008; Ranjith-Kumar et al., 2001; Tomar et al., 2006), which is comparable to cellular TNTases (Martin and Keller, 2007). What remains to be elucidated is how the reaction happens: during RNA synthesis, the 3' end of the template RNA is positioned in the template channel of the RdRp, which adds nucleotides onto the 3' end of the nascent RNA, positioned in the active site. So, to add non-templated nucleotides to the 3' end either the RdRp turns into the opposite orientation

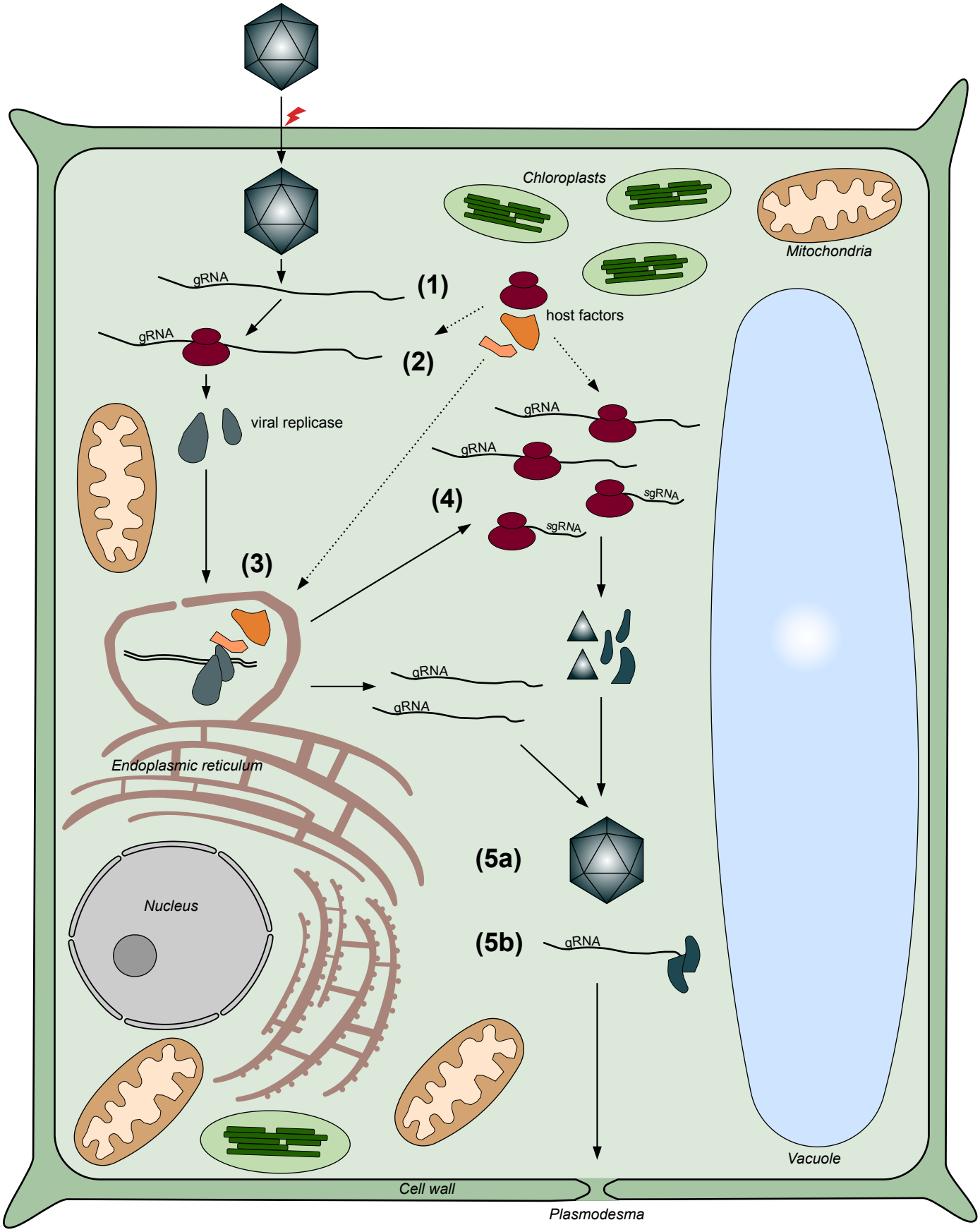


Figure 7: Overview of the viral cycle of a ss(+)RNA phytovirus. (1) After penetrating into the cell, the genome is disencapsidated. **(2)** The first viral proteins are translated using the host translation complex. **(3)** In organelle membrane-derived vesicles, the viral genome is replicated by viral proteins in cooperation with host factors. **(4)** The other viral proteins are translated from gRNA or sgRNAs. **(5a)** *De novo* synthetized viral genomes are packed by the capsid protein and new virions can move into another cell. **(5b)** The propagation of the viral infection in other cells can also be perpetrated by the movement of the viral RNA, helped by viral and host proteins.

and pulls the 3' extremity of the template RNA into its active site (Poranen et al., 2008), or it could further add nucleotides to the nascent RNA, without relating to the template RNA. Moreover, even if the residues involved in the nucleotide specificity are identified, why some viral RdRp evolved towards this nucleotide specificity remains to be understood.

The addition of non-templated nucleotides to the 3' end of viral RNAs could also be performed by cellular enzymes. In line with this hypothesis, the poly(A) tail on HCV RNAs of some infectious clones is followed by a guanosine (van Leeuwen et al., 2006). Guanosines in viral poly(A) tails have also been recently found on hepatitis B virus mRNA and on human cytomegalovirus RNA 2.7. These mixed tails have been shown to stabilize viral RNAs. In this study, the authors also describe that these tails are added by a cellular TNTase, TENT4, that is recruited to the viral RNA on a stem-loop in the 3'UTR, via an adaptor protein (Kim et al., 2020a).

c. TNTase activity in viral genomes

3' non-templated nucleotide additions to viral RNAs remains rarely studied. Studies in the 1990s-2000s and a recent one on coronaviruses identified terminal nucleotidyl transferase activities of viral RdRps (Arnold et al., 1999; Behrens et al., 1996; Fullerton et al., 2007; Neufeld et al., 1994; Ranjith-Kumar et al., 2001; Rohayem et al., 2006; Smallwood and Moyer, 1993; Tomar et al., 2006; Tvarogová et al., 2019; Wang et al., 2013b; Wu et al., 2014). TNTase activity of a viral RdRp has first been described for the *Vesicular stomatitis virus* (Smallwood and Moyer, 1993). The following year, a terminal adenylyl transferase activity has been reported for the poliovirus RdRp. Interestingly, the authors report that the viral RdRp activity is specific for adenosines and uridines, and that the adenosines are added by block of three (Neufeld et al., 1994). This activity could be involved in the poly(A) tail repair. The RdRp of the Sindbis virus (Tomar et al., 2006) and of the human coronavirus 229E (Tvarogová et al., 2019) have also been reported to possess a TNTase activity, with preference for adenosines, suggesting a role in the synthesis and repair of the poly(A) tail. Other viral RdRp have been described to have a preference for uridines, like the RdRps of the poliovirus, hepatitis C virus, Wuhan nodavirus and flock house virus (Arnold et al., 1999; Behrens et al., 1996; Wang et al., 2013b; Wu et al., 2014).

Uridines and adenosines are not the only substrates of the TNTase activity of viral RdRps: for two species of *Flaviviridae*, the hepatitis C virus and the bovine viral diarrhea virus, the RdRp can add non-templated cytosines, adenosines or uridines according to the recognized 3' terminal nucleotide. The authors indeed show that in the case of a 3' cytosine or guanosine, it will preferentially add a cytosine, whereas if the 3' terminal nucleotide is an adenosine, the RdRp will further add adenosines (Ranjith-Kumar et al., 2001). Another study demonstrates that, *in vitro*, the purified RdRp of a *Sapovirus* (*Caliciviridae*) can add non-templated nucleotides, with a preference for cytosines (Fullerton et al., 2007).

Collectively, these data indicate that various activities can generate this non-templated nucleotide additions. Cellular-encoded TNTase activities play an important role in RNA metabolism (De Almeida et al., 2018b; Warkocki et al., 2018a; Zigáčková and Vaňáčová, 2018), so their involvement in viral RNA 3' tailing would not be surprising.

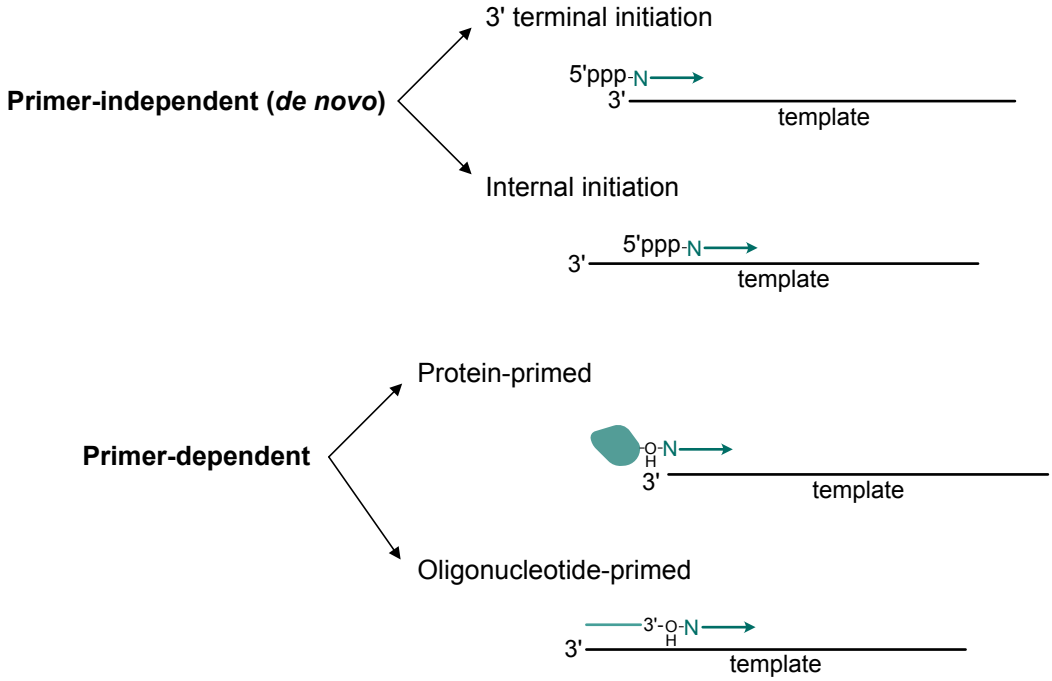


Figure 8: Strategies of initiation of the genome replication for (+) RNA viruses. There are two mechanisms of RNA synthesis by (+) RNA virus RNA polymerases: *de novo* initiation or primer-dependent initiation. *De novo* initiation either starts from the 3' terminal nucleotide or at an internal site. In primed-initiation, the primer is either a protein, like a VPg (green) bound to nucleotide(s) (N, orange), or an oligonucleotide. A hydroxyl group of a tyrosine or serine residue of a protein or the 3' hydroxyl group of a nucleotide provides the hydroxyl group required for the formation of the phosphodiester bond with the first nucleotide. Adapted from Paul and Wimmer, 2015.

3. Interaction between ss(+) RNA viruses and plants

In this chapter, I will present the viral cycle and then focus on plant defenses against viral infection.

1. The viral cycle of ss(+) RNA viruses in plants

RNA viruses with a genome of positive polarity replicate in the cytoplasm (Fig. 7). Once it has entered the plant cell, the viral genome is liberated into the cytoplasm. The second step is the translation of essential proteins such as the viral RdRp, to produce new viral genomes. Then, an equilibrium between replication and translation ensures the production of viral genomes and viral proteins, that will be assembled into new virions. To prevent the recognition of double-stranded RNA intermediates generated during viral replication by the host surveillance pathway and the subsequent RNA interference response (or interferon response in animals), the viral replication occurs in cell membrane-derived vesicles and in plants, most viral genomes encode for proteins with silencing suppressor activity (Bellott et al., 2019).

a. Genome replication

There are two strategies to initiate the viral RNA genome replication (Fig. 8): *de novo* or primed-initiation. In *de novo* initiation, the viral RdRp either starts at the last 3' terminal nucleotide or at an internal site (Beerens et al., 2007; Kao and Ahlquist, 1992) and this strategy is mostly used by ss(+) RNA viruses (Paul and Wimmer, 2015). Internal initiation of replication is reported in plant viruses for the cucumber necrosis virus (*Tombusvirus*, (Panavas et al., 2002)) and TYMV and BMV, as previously mentioned in the "TLS" part (Chapman and Kao, 1999; Deiman et al., 1998; Dreher and Hall, 1988; Dreher et al., 1984; Singh and Dreher, 1998).

Primer-dependent initiation is either protein-primed or oligonucleotide-primed (Nagy et al., 1997; Paul et al., 1998). The protein-primed initiation mechanism is used by picornaviruses, in particular the poliovirus. Poliovirus genome replication mechanism has been extensively studied (Paul and Wimmer, 2015). Both RNA strands have a 5' VPg (Ambros and Baltimore, 1978; Lee et al., 1977; Rothberg et al., 1978). To initiate the replication, the VPg is uridylylated, *i.e.* two uridines residues are covalently linked to one of its amino acids, generating a VPg-pUpU form. VPg nucleotidylation is crucial to incorporate the VPg onto the viral genome. The free hydroxyl group required for the nucleotidylation reaction is provided either by a tyrosine, a threonine or a serine residue present in the VPg. For example, a tyrosine residue in the VPgs from *Picornaviridae*, *Potyviridae* and *Secoviridae* viral species provides the hydroxyl group. In sobemoviruses, a threonine or a serine residue is used (Ambros and Baltimore, 1978; Olspert et al., 2011a; Pinck et al., 1991). Of note, the role of the VPg in the initiation of the replication is not confirmed for species from the *Sobemoviridae* family (Sõmera et al., 2015) nor from the *Secoviridae* family, in particular *Nepovirus* species (Fuchs et al., 2017; Thompson et al., 2017).

In plants, the tyrosine residue was reported to be essential for viral infection of some potyviruses (Murphy et al., 1996). Later on, two studies have shown for the potato virus A and the pepper vein banding virus that the VPg is nucleotidylated on the tyrosine residue (Anindya et al., 2005; Puustinen and Mäkinen, 2004).

b. Translation of the viral genome

Viral RNAs of positive polarity are directly competent for translation. For the plant viral RNA genomes that have a 5' cap and a 3' poly(A), translation is initiated like for cellular mRNAs (Fütterer and Hohn, 1996). If there is no cap, cap-independent translation relies on internal binding sequences within the 5' and the 3' UTRs of the viral genome: the internal ribosome entry site (IRES) at the 5' extremity and the 3' cap-independent translation enhancer (3'CITE). IRESes were first discovered for an animal virus (Pelletier et al., 1988) but since then, many articles report these structures for plant viruses as well. The plant viral IRESes are shorter and less structured than the animal ones. Most studies focus on viral species from the *Potyviridae* family but the carmovirus TCV RNA also possesses an IRES. IRES allows the binding of translation factors such as eIF4G or eIF4E, as well as ribosomal RNA (Geng et al., 2021). In the 3'UTR, 3'CITEs were first characterized for plant viruses (Danthinne et al., 1993). There are now many examples of 3'CITEs described in plants and they were recently reviewed (Geng et al., 2021).

Another important difference between viral RNAs and cellular mRNAs is that viral RNAs are usually polycistronic. Some polycistronic viral RNAs can be translated into a polyprotein that is next cleaved by viral protease activity, like it is found in *Potyviridae* species (Revers and García, 2015). Viruses evolved towards several, not mutually exclusive, strategies to translate all their open reading frames (ORFs). The segmentation of the genome allows to present the initiation codons from an initially polycistronic genomic RNA to the translation complex. In some cases, subgenomic RNAs are produced by the RdRp. During RNA replication, a complementary RNA (cRNA) is synthetized using the genomic RNA as template. This cRNA is then used by the RdRp as template to produce shorter RNAs, the sgRNAs, that each possesses the same 5' and 3' characteristics as the genomic RNA. For example, the potato virus X (PVX, *Alphaflexviridae*) genomic and subgenomic RNAs have a 5' cap and a 3' poly(A) tail (Verchot, 2022).

During the translation, viruses also use other tricks to further express viral proteins: leaky scanning, non-AUG initiation, ribosomal frameshift and ribosomal read-through (Geng et al., 2021). The leaky scanning is a mechanism broadly used by plant viruses. An ORF located within another ORF can be expressed if the ribosome skips the first initiation codon and starts to translate at the downstream initiation codon, due to the non-optimal context of the first AUG (Kozak, 2002). Examples of leaky scanning have been reported in members of the *Virgaviridae* and *Tombusviridae* families (Castaño et al., 2009; Herzog et al., 1995; Turina et al., 2000).

Ribosomal frameshift allows the translating ribosome to change the reading frame of the ORF and has initially been described in an animal virus (Hizi et al., 1987). One or two nucleotides to the 5' or to the 3' direction are skipped by the ribosome (Miller and Giedroc, 2010; Penn et al., 2020). This mechanism has been described for several plant viruses, as for example in the translation of the PIPO protein from the *Potyviridae* TuMV (Chung et al., 2008).

Another way to express ORFs from a polycistronic RNA is ribosomal readthrough during which the ribosome misses the stop codon and moves through. This was initially reported for the *Virgaviridae* TMV (Beier et al., 1984b, 1984a) and then for other viral species from the *Benyvirus* genus and from the *Solemoviridae*, *Virgaviridae*, *Tombusviridae* and the *Tymoviridae* families (Beier and Grimm, 2001;

Malpica-López et al., 2018). Large viral genomes are often expressed via a combination of these strategies (Agranovsky et al., 1994).

2. Host defense against ss(+) RNA viruses in plants

a. The plant immune response to ss(+) RNA viruses

Plants cells are protected from outside aggressions by a thick cell wall topped by a waxy cuticular layer. If these barriers are breached, the pathogen will face an active plant immune system that is able to specifically recognize it and to trigger a series of responses. To defend against viral infection, plants evolved towards several non-mutually exclusive strategies: an innate immune response, RNA silencing, translational repression and protein degradation (Wu et al., 2019).

Pathogen-triggered immunity (PTI)

Like metazoans, plants have pattern recognition receptors (PRRs) that are localized outside the cell (DeFalco and Zipfel, 2021; Monaghan and Zipfel, 2012) and that can recognize pathogen- or microbe-associated molecular patterns (PAMPs/MAMPs) (DeFalco and Zipfel, 2021). These proteins resemble the Toll-like receptors in animals, as they also are kinases containing leucine-rich repeats and lysine motifs. Once the PAMP has been recognized, it results in the propagation of an intracellular signaling that leads to transcriptional reprogramming and biosynthesis of the factors involved in the immune response (DeFalco and Zipfel, 2021; Monaghan and Zipfel, 2012).

The PAMP recognition outside of the cell by the PRR is the first perception of the pathogen by a plant and it induces the pathogen-triggered immune response (PTI). The PRR dimerizes and binds to other proteins to trigger downstream intracellular signaling, including the biosynthesis of defense hormones, the synthesis and deposition of callose at plasmodesmata and the expression of pathogen-related genes (Couto and Zipfel, 2016; Tang et al., 2017). Sometimes, the hypersensitive response (HR) is activated, resulting in cell death and visible by necrotic spots on leaves (Coll et al., 2011). To date, plant PAMPs and PRRs are largely described in the case of bacterial infections (Boutrot and Zipfel, 2017), but little is known about the pathogen-induced immunity against viruses. The mutations of some PRRs have been reported to increase the sensibility to RNA viruses (KØrner et al., 2013; Yang et al., 2010). A PTI-like response can be induced by viral coat proteins. TMV coat protein detection outside tobacco cells triggers the production of reactive oxygen species (Allan et al., 2001). A calcium-depend signal cascade leading to the inhibition of cell-to-cell movement has been observed upon the exposition to the TGBp1 and the coat proteins of PVX (Perraki et al., 2018).

Effector-triggered immunity (ETI)

During the arms race of evolution between hosts and pathogens, the latter also have evolved factors that are able to compromise the PTI defenses of the host. In this case, plants have another system relying on intracellular proteins that can recognize these pathogen effectors. The effector-triggered immunity (ETI) is guaranteed by R genes that code for proteins containing a N-terminal Toll/Interleukin-1 homology domain or a coiled-coil domain, a nucleotide-binding domain in the middle and a leucine-rich repeat domain at the C-terminal extremity (Chisholm et al., 2006). The R genes-mediated response mainly results in HR (Cui et al., 2015). Of note, both PTI and ETI can lead to systemic

acquired resistance (Fu and Dong, 2013). Other genes independent from the classical R genes have been described to confer viral resistance (Wu et al., 2019).

Virus mechanisms to counteract plant immune response

Viruses are not powerless against plant immunity, as some viral proteins can alter innate immunity like the capsid protein of the plum pox virus (*Potyviridae*), the RdRp of TuMV and the movement protein of CMV (Cheng et al., 2017; Kong et al., 2018; Nicaise and Candresse, 2017) but the underlying molecular mechanism remains unclear for PPV and CMV proteins.

Although the molecular mechanisms remain unclear, ribosome-inactivating proteins have been reported to be able to have an antiviral action on several plant viruses including AlMV, BMV, CMV, PVX, TBSV, TMV, TEV and TuMV (Citores et al., 2021; Domashevskiy et al., 2017). Arabidopsis recovery from TRV infection was reported to involve translational repression (Ma et al., 2015). In addition to translation repression, the viral proteins levels are controlled by protein degradation via the ubiquitination and the autophagy pathways (Wu et al., 2019). Examples of viruses counteracting viral protein degradation have been described, like PVX (Chiu et al., 2010), TYMV (Chenon et al., 2012), species from the *Solemoviridae* (Baumberger et al., 2007; Derrien et al., 2012) and TuMV (Cheng and Wang, 2016; Hafrén et al., 2018).

b. RNA silencing: a balance between decay and translation repression

One of the most extensively studied antiviral pathway in plants is probably RNA silencing or RNA interference. This evolutionary conserved mechanism has been extensively studied for its role in plant development, the regulation of transposable elements and the antiviral response (Baulcombe, 2004; Ding and Voinnet, 2007). Classically, a dsRNA molecule is recognized and processed by DICER-LIKE (DCL) proteins, that are type III endonucleases expressed in plants. DCL proteins dice the dsRNA into small 20-24 nt RNA duplexes that are loaded onto ARGONAUTE (AGO) proteins, the core of the RNA-induced silencing complex (RISC).

This complex can cleave viral RNAs and/or suppress viral protein translation in a sequence specific manner (Llave, 2010; Pantaleo et al., 2007). The viral RNA cleavage was reported to occur in hot spots, preferentially on the minus strand (Pantaleo et al., 2007). The RNA silencing mechanism can be amplified by a process called transitivity, during which secondary small interfering RNAs (siRNAs) targeting the viral RNA are produced (de Felippes and Waterhouse, 2020; Himber et al., 2003). These siRNAs are able to move throughout the whole plant (Molnar et al., 2011).

To avoid the activation of RNA silencing through the detection of dsRNA molecules, the replication of the viral genome takes place in membrane-derived vesicles formed by the action of viral proteins, also called viral factories (Hyodo and Okuno, 2014; Jin et al., 2018). The membranes of all the organelles can be used (Xu and Nagy, 2014). Two viral species from the *Tombusviridae*, the carnation Italian ringspot virus and the tomato bushy stunt virus use the mitochondrial membrane and the peroxisome membrane, respectively (Barajas et al., 2009; Hwang et al., 2008). BMV, TMV, TEV and TuMV use the endoplasmic reticulum membrane (Cotton et al., 2009; Más and Beachy, 1999; Restrepo-Hartwig and Ahlquist, 1996; Wei and Wang, 2008). Viral replication on chloroplast membranes has been described for a hordeivirus (Carroll, 1970; Torrance et al., 2006), potyviruses, including TEV and TuMV (Gadh and

Hari, 1986; Kitajima and Costa, 1973; Martin et al., 1995; Mayhew and Ford, 1974) and the tymovirus TYMV (Lafleche et al., 1972). These data indicate that some plant viruses display a certain flexibility in the selection of subcellular compartments for viral factories.

In addition, plant viruses express proteins that can disrupt RNA silencing (Bellott et al., 2019). These viral silencing suppressors (VSRs) do not share amino acids or structure resemblances, suggesting they evolved independently. Functionally, there are two types of VSRs: either they act on the virus-derived siRNAs such as the protein p19 in tombusviruses or the p122 from tobamoviruses, or they can target RNA silencing factors like the VPg of potyviruses, the TGBp1 of potexviruses or the P0 of poleroviruses (Bellott et al., 2019). The HC-Pro of potyviruses can act at both levels (Bellott et al., 2019; Hu et al., 2020; Valli et al., 2018). Moreover, the p19 protein of a tombusvirus has been shown to upregulate the production of miR168 that target *AGO1* transcripts, resulting in the downregulation of AGO1 (Bellott et al., 2019; Várallyay et al., 2010).

RNA silencing can also mediate translation repression of viral proteins (Brodersen et al., 2008; Valencia-Sanchez et al., 2006). AGO1-mediated translation repression of ToRSV RNA2 is linked to recovery in tobacco (Ghoshal and Sanfaçon, 2014).

Of course, these antiviral mechanisms are interconnected. They can depend on or regulate each other, and they can cooperate against pathogens (Wu et al., 2019).

	S. cerevisiae	*S. pombe*	*C. elegans*	*D. melanogaster*	*H. sapiens*	*A. thaliana*
Decapping complex and decapping activators	Dcp1	Dcp1	Dcp1	Dcp1	DCP1	DCP1
	Dcp2	Dcp2	Dcp2	Dcp2	DCP2	DCP2
	Lsm1-7	Lsm1-7	Lsm1-7	Lsm1-7	LSM1-7	LSM1-7
	Edc1, Edc2, Edc3	Edc1, Edc3	-	Edc3	EDC3	-
	Pat1	Pat1/Pdc2	PATR-1	Hpat/Patr-1	PATL1	PAT1, PATH1, PATH2
	Scd6	Scd6	CAR-1	Trailer Hitch (Tra1)	LSM14A/SCD6	DCP5, DCP5-L
	Dhh1	Dhh1/Ste13	CGH-1	Me31b	DDX6	RH6/8/12
	Pby1	-	-	-	-	-
	-	-	4E-T/Cup	Cup	4E-T/Cup	-
	-	Pdc1	Edc4	Edc4/Ge-1	EDC4/HEDLS	VCS, VCR
Deadenylation	-	PARN	PARN	-	PARN	PARN
	Pan2/Pan3	Pan2/Pan3	Pan2/Pan3	Pan2/Pan3	PAN2/PAN3	-
	Ccr4-Not: Ccr4, Pop2	Ccr4-Not: Ccr4, Caf1	Ccr4-Not	Ccr4-Not: Ccr4, Pop2	CCR4-NOT: CCR4, CAF1	CCR4-NOT: CCR4 (2 paralogs), CAF1 (11 paralogs)
5'-3' exoribonuclease	Xrn1	Exo2	Xrn1	Pacman	XRN1	XRN4
3'-5' exoribonuclease	Dis3L2	Dis3L2	Dis3l-2	Dis3l2	DIS3L2	SOV
RNA exosome (core)	Exo9	Exo9	Exo9	Exo9	Exo9	Exo9
Exosome 3'-5' exo/endo-ribonucleolytic activities	Dis3/Rrp44	Dis3/Rrp44	Dis3/Rrp44	Dis3/Rrp44	DIS3/RRP44	RRP44
	-	-	-	-	DIS3L1	-
Ski complex	Ski2, Ski3, Ski7, Ski8	Ski2, Ski3, Ski7, Ski8	Ski2, Ski3, Ski7, Ski8	Ski2, Ski3, Ski7, Ski8	SKI2, SKI3, SKI7, SKI8	SKI2, SKI3, SKI7, SKI8
Other exosome cofactors	SKA1	-	-	-	FOCAD, AVEN	RST1, RIPR
TUTases	-	Cid1	CDE-1	Tailor	TUT4, TUT7	URT1, HESO1

Table 1: Factors identified to be involved in cytosolic RNA decay in eukaryotes.

4. RNA degradation pathways in eukaryotes

RNA degradation is important to regulate gene expression in all organisms. RNA degradation is crucial for cell homeostasis and to remodel the transcriptome upon stresses or developmental transitions. In addition, RNA decay participates to fight viral infections by targeting viral RNAs. RNAs can be degraded either from their 5' or their 3' terminus or by endoribonucleolytic cleavage, followed by the degradation of the resulting fragments. mRNA degradation in the cytoplasm involves many players (Table 1, Fig. 9) and usually starts by the elimination of their terminal features, first the poly(A) tail and then the cap structure.

1. Factors and mechanisms involved in bulk RNA decay

a. The deadenylases

The limiting step of mRNA half-life lies in the shortening of the poly(A) tail by deadenylases. Deadenylation is essential to trigger both 5'-3' and 3'-5' degradation pathways, if no endoribonucleolytic cleavage has occurred on the RNA (Fig. 9). Three enzymes/complexes have been characterized in eukaryotes so far: the poly(A) ribonuclease PARN, the poly(A)-specific ribonuclease complex PAN2/3, and the Ccr4/Not complex.

The poly(A) ribonuclease PARN

The poly(A) ribonuclease PARN is a 3'-5' exoribonuclease conserved in eukaryotes, except for *S. cerevisiae* and *D. melanogaster* (Parker and Song, 2004). Two PARN interacting as homodimer bind to the poly(A) tail via their RRM domain (Wu et al., 2005). Moreover, if the m^7G cap structure is accessible and not bound by the cap binding complex, PARN can directly associates with the m^7G cap structure, which regulates its activity of deadenylation (Balatsos et al., 2006; Dehlin et al., 2000; Martînez et al., 2001). PARN functions in xenopus oocytes early development (Copeland and Wormington, 2001). In mammalian cells, this deadenylase also targets mRNAs containing AU-rich elements in cell extracts (Lai et al., 2003), RISC-cleaved transcripts (Zhang et al., 2015a), as well as NMD targets (Lejeune et al., 2003). In Arabidopsis, mutations in *PARN* are embryo lethal (Chiba et al., 2004; Reverdatto et al., 2004) and PARN was proposed to target a subset of mRNAs involved during embryogenesis (Reverdatto et al., 2004). PARN was first reported to be localized within the nucleus and the cytosol (Chiba et al., 2004; Reverdatto et al., 2004). Yet, those localizations were determined using N-terminal GFP fusions. Conversely, by using C-terminal GFP fusions, PARN was later shown to localize within mitochondria (Hirayama et al., 2013). These authors further show that PARN, together with a mitochondrial PAP, regulate the polyadenylation status of mitochondrial mRNAs (Hirayama et al., 2013). This regulation has been proposed to be conserved within land plants. Indeed, plant PARNs possess a N-terminal extension that is not found in animal PARNs and could explain their mitochondrial localization (Hirayama, 2014; Kanazawa et al., 2020).

The poly(A)-specific ribonuclease complex PAN2/PAN3

In yeast, the heterotrimeric complex Pan2/Pan3 is composed of one Pan2 and two Pan3 proteins (Schäfer et al., 2019; Wolf et al., 2014). It is recruited on the poly(A) tail through the interfaces formed by PABP oligomerization on the poly(A) tail and its processivity is contingent on the number of bound PABPs (Schäfer et al., 2019). No homolog of *PAN2* is found in higher plants, although it is conserved in Chlorophytes and Bryophytes (Pavlopoulou et al., 2013).

The Ccr4/Not complex

The Ccr4/Not complex is thought to act as the main deadenylase activity in eukaryotes. This activity is conferred by two catalytic subunits, Ccr4 and Caf1. Apart from deadenylation, the Ccr4/Not complex is involved in many different pathways such as chromatin modifications, transcription regulation, mRNA quality control, nuclear export in the nucleus, as well as translation regulation and protein quality control in the cytoplasm (Collart, 2016). To ensure deadenylation, it is recruited by RNA binding proteins such as PUF and TTP (Webster et al., 2019) in fission yeast, the RISC complex (Behm-Ansmant et al., 2006;

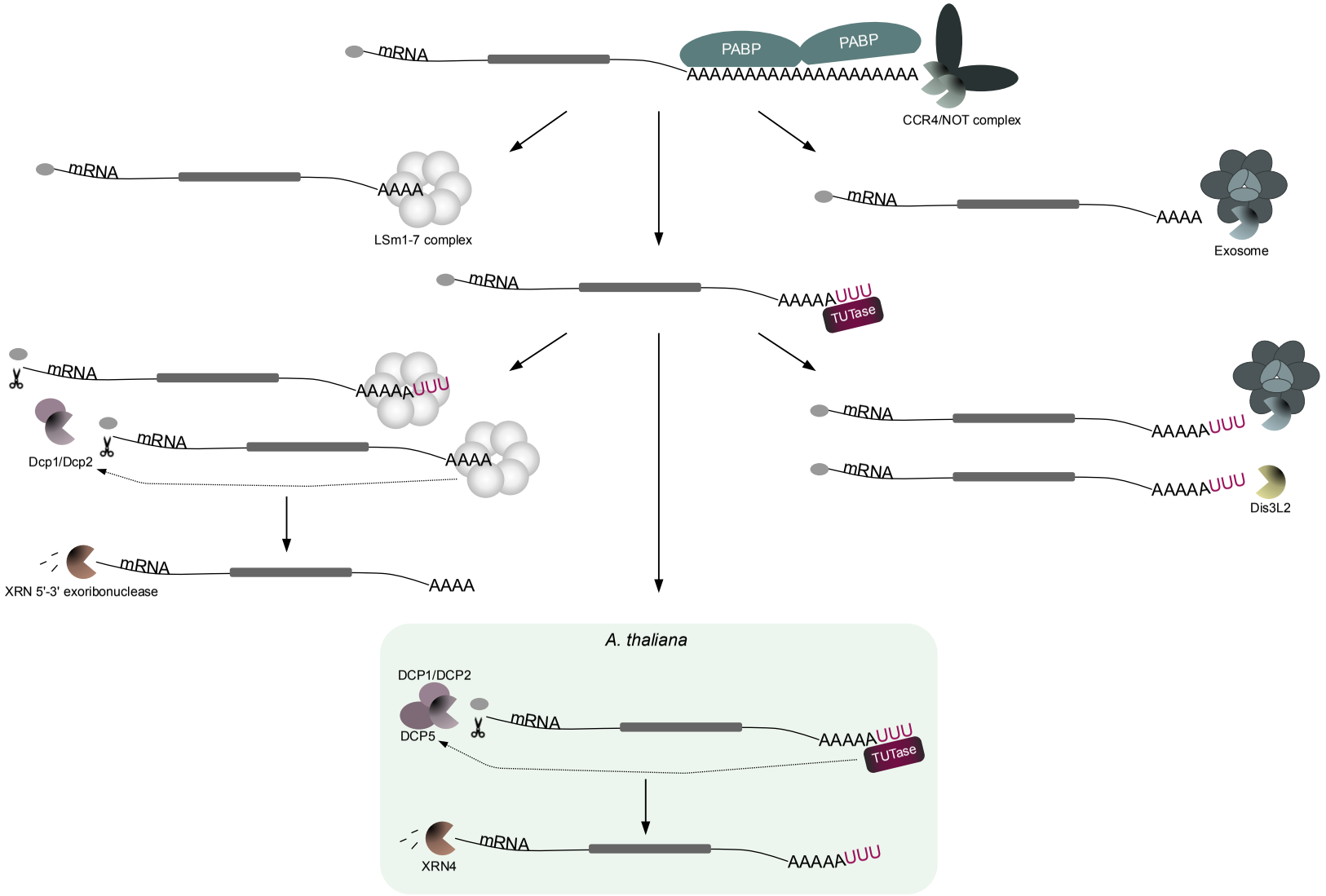


Figure 9: Overview of cytoplasmic mRNA decay pathways in eukaryotes. After a deadenylation step performed by the CCR4/NOT complex, the deadenylated mRNA can be further degraded from its 3' end by the RNA exosome. Mostly, the oligo(A) is recognized by the LSm1-7 complex that recruits the decapping complex. Deadenylated mRNAs can also be uridylated, which further enhances the binding of LSm1-7 and subsequent decapping. After the cleavage of the 5' cap structure, the accessible 5' phosphate is a substrate for the 5'-3' exoribonucleases XRN. 3' uridylation also stimulates the binding of the 3'-5' exoribonuclease Dis3L2 or the exosome, and subsequent degradation from the 3' extremity of the mRNA. In *A. thaliana*, uridylation stimulates the decapping via the interaction between the TUTase and a decapping activator. The mRNA is then degraded from its 5' extremity.

Fabian et al., 2011) or via YTHDF2 proteins on RNA modifications like m^6A (Du et al., 2016) in mammalian cells. Recently, Ccr4/Not interaction with the ribosome has been reported in the case of low optimal codon translation, which helps to understand the link between low codon optimality and mRNA degradation (Buschauer et al., 2020). All subunits of the complex have homologs in plants but its biological functions are still largely unknown. It is interesting to note that the *CAF1* family is largely expanded in Angiosperms, with 12 CAF-encoding genes (Pavlopoulou et al., 2013; Walley et al., 2010) and two *CCR4* paralogs, *CCR4a* and *CCR4b* (Dupressoir et al., 2001).

To shorten the poly(A) tail in human cell lines, it has originally been proposed that the PAN2/3 deadenylases act prior to the Ccr4/Not complex (Yamashita et al., 2005). PAN3 recruits PAN2 to the mRNA 3' terminus via the PABPCs (Uchida et al., 2004) and PAN2 reduces the poly(A) tail length in a distributive manner (Yamashita et al., 2005). Ccr4 completes the deadenylation, first in a processive manner and when the tail reaches 45 nt, it switches to distributive mode (Yamashita et al., 2005). Yet, recent studies in yeast led to reconsider this model of sequential deadenylation. Degradation rate profile analysis in yeast depleted from Pan2, Pan3, Ccr4, Pop2 suggest that Pan2/3 and Ccr4/Not complexes may target different transcripts subpopulations (Sun et al., 2013). Both can initiate deadenylation, but only the Ccr4/Not complex is able to complete it and to release the last PABP. In addition, there are functional differences within the Ccr4/Not complex. Caf1 preferentially targets transcripts with a low codon optimality and in a translation-dependent manner, while Ccr4 targets all mRNAs (Webster et al., 2018).

As PAN2/3 are not conserved across the plant lineage (Pavlopoulou et al., 2013), the Ccr4/Not complex could ensure the whole deadenylation process in plants, from its initiation to its ending.

Nocturnin and HESPERIN

Of note, a fourth deadenylase activity has been proposed for Nocturnin (NOCT) in metazoans. This protein shares similarities with CCR4 (Godwin et al., 2013). Its expression in metazoans is controlled by the circadian clock and reaches its highest at night, as its name reminds it (Green and Besharse, 1996; Li et al., 2008). NOC function as a deadenylase was proposed on the basis of *in vitro* assays on crude NOCT preparations (Baggs and Green, 2003). Yet, several studies reported the lack of deadenylase activity for purified NOCT *in vitro* (Abshire et al., 2018; Estrella et al., 2018). In line with this observation, a recent study reported that NOCT is targeted to the mitochondria and identified the dinucleotides $NADP^+$ and NADPH as NOCT substrates in drosophila and animal cells (Estrella et al., 2019). The ortholog of NOC in Arabidopsis, HESPERIN (HESP), also influences the steady-state level of circadian transcripts (Delis et al., 2016), but its deadenylase activity remains to be confirmed using a catalytic mutant and adequate RNA substrates.

b. The 5'-3' decay pathway: interplay between deadenylation and decapping

The decapping complex and its activators

After mRNA deadenylation, the 5' cap structure is removed, followed by the mRNA degradation from its 5' extremity, accessible to exoribonucleases (Fig. 9). Dcp2 is the main decapping enzyme in eukaryotes. First described in yeast (Dunckley and Parker, 1999), this protein from the Nudix hydrolase family is conserved in Eukaryotes (Cohen et al., 2005; Van Dijk et al., 2002; Iwasaki et al., 2007; Lin et

al., 2008; Lykke-Andersen, 2002; Wang et al., 2002; Xu et al., 2006). Dcp2 is composed of a N-terminal regulatory domain, a catalytic core constituted of a Box A domain and a Nudix hydrolase domain, and an intrinsically disordered C-terminal region (Vidya and Duchaine, 2022). The Nudix hydrolase domain confers the activity to specifically cleave the alpha-beta pyrophosphate bond of capped mRNAs (Van Dijk et al., 2002; Lykke-Andersen, 2002; Wang et al., 2002). The C-terminal intrinsically disordered region is common in all eukaryotic forms of Dcp2 and is mostly studied in yeast. It has been reported to contain positive or negative regulatory elements of Dcp2 activity (Vidya and Duchaine, 2022). The decapping activity is ensured by the Dcp1/Dcp2 holoenzyme. Dcp1 is the main regulator of Dcp2 activity (Sakuno et al., 2004) and interacts with Dcp2 N-terminal domain via the EVH1 domain (She et al., 2004, 2008).

Other decapping activators are described, such as the Lsm1-7 complex. This complex, conserved in Eukaryotes, binds to oligoadenylated RNA and enhances decapping activity (Chowdhury et al., 2007). In budding yeast *S. cerevisiae*, other decapping activators are Edc1-3, Pat1, Scd6 and the RNA helicase Dhh1 (Nissan et al., 2010). In fission yeast *S. pombe*, the decapping activators Edc3, Scd6, Ste13 (homolog to the RNA helicase Dhh1, (Fromm et al., 2012)) and Pdc1 (a WD40-like protein, (Wang et al., 2013a)) were characterized. Fission yeast Pdc2 can trigger mRNA decay by two different mechanisms, either by interacting with the Lsm1-7 complex or by directly binding to Dcp2 (Lobel et al., 2019). In *D. melanogaster*, the decapping activators described are Edc3, Tral (homolog to Scd6), Hpat (homolog to Pat1, (Haas et al., 2010)) and the RNA helicase Me31b (homolog to Dhh1, (Tritschler et al., 2008)). In human cells, the interaction between DCP1 and DCP2 is mediated through EDC4, sharing similarities with *S. pombe* Pdc1 (Wang et al., 2013a). Other decapping factors in humans involve EDC1-3, LSM14A/SCD6 and the RNA helicase DDX6 (Vidya and Duchaine, 2022).

In plants, DCP1 and DCP2 interact with a third protein, VARICOSE (VCS), homolog to human EDC4 (Deyholos et al., 2003). DCP2 activity is stimulated by DCP1 and VCS (Xu et al., 2006). Other decapping activators involve the RNA helicases RH6/8/12 (homologs to Dhh1, (Chantarachot et al., 2020; Xu et al., 2006)), PAT1 (Roux et al., 2015) and the LSM domain-containing protein DCP5, homolog of human LSM14 and yeast Scd6 (Xu and Chua, 2009). The LSM1-7 complex also participates in the 5'-3' degradation process in plants. Two studies reported that capped mRNAs are stabilized in a *lsm1a lsm1b* mutant (Golisz et al., 2013; Perea-Resa et al., 2013). Further work is however required to elucidate the whole role of the LSM1-7 complex in plant cytosolic RNA decay.

The 5'-3' exoribonucleases of the XRN family

After the removal of the cap structure, the mRNA 5' extremity is accessible for degradation by 5'-3' exoribonucleases of the eXoRiboNuclease (XRN) family. All XRN proteins have two conserved regions CR1 and CR2. In the CR1 region, there is a seven amino acids motif that is strictly evolutionary conserved and is involved in the Mg^{2+} ion coordination during catalysis (Yang et al., 2006a). The active site is structured so that XRN proteins specifically target single strand RNA with an exposed 5' phosphate, thus m^7G or triphosphorylated RNAs do not fit in the pocket. In addition, XRN substrates must have a sufficient length of single-stranded RNA, 4 nt or more, to attain the active site (Jinek et al., 2011).

In yeast, Xrn1 and Xrn2 are the main 5'-3' exoribonucleases (Amberg et al., 1992; Stevens, 1978). Orthologs of these proteins were discovered in all key model organisms: *S. pombe* (Käslin and Heyer, 1994; Shobuike et al., 2001), *M. musculus* (Bashkirov et al., 1997; Shobuike et al., 1995), *H. sapiens* (Sato et al., 1998; West et al., 2004), *D. melanogaster* (Till et al., 1998), *A. thaliana* (Kastenmayer and Green, 2000), *C. elegans* (Chatterjee and Großhans, 2009; Newbury and Woollard, 2004) and *T. brucei* (Li et al., 2006). Interestingly, there is no Xrn1 ortholog in higher plants, but three Xrn2 homologs have been identified in the Arabidopsis genome (Kastenmayer and Green, 2000). Although it is a homolog of Xrn2, XRN4 fails to complement the *xrn2* mutation in yeast. Instead, it is able to complement a *xrn1* mutation (Kastenmayer et al., 2001), indicating that XRN4 is a structural homolog of yeast Xrn2 but a functional homolog of yeast Xrn1.

In yeast, Xrn1 has been reported to interact with the decapping proteins Dcp1 and Dcp2 (Bouveret et al., 2000; Fromont-Racine et al., 2000; Tharun et al., 2000). Xrn1 also interacts with the Lsm1-7 complex (Bouveret et al., 2000) and with Pat1 (Nissan et al., 2010). In mammalian cells, NMD factors Upf proteins have been reported to interact with Xrn1 (Lejeune et al., 2003).

XRN proteins are localized differently within the cell. Xrn1 (in plants XRN4) has been described to localize within the cytoplasm, predominantly in foci formed by liquid-liquid phase separation called processing bodies (P-bodies). Xrn1 has also been detected in other cytoplasmic granules like stress granules, neuronal granules or germ cell granules. Xrn2 (in plants XRN2 and XRN3) on the other hand is found within the nucleus and the nucleolus (Nagarajan et al., 2013).

XRN proteins are involved in different molecular mechanisms, but the best known role of Xrn1 (in plants XRN4) is to degrade uncapped mRNAs (Gazzani et al., 2004; Hsu and Stevens, 1993; Lejeune et al., 2003; Muhlrad et al., 1994; Rymarquis et al., 2011), confirmed by the interaction with the decapping machinery in yeast (Nissan et al., 2010). 3' fragments generated from mRNA endonucleolytic cleavage also are degraded by Xrn1 (XRN4) in *S. cerevisiae*, human cells, *D. melanogaster* and *A. thaliana* (Doma and Parker, 2006; Eberle et al., 2009; Gatfield and Izaurralde, 2004; Nagarajan et al., 2019; Orban and Izaurralde, 2005; Souret et al., 2004). A recent transcriptome-wide study in Arabidopsis has reported that XRN4 is involved in the decay of many types of transcripts, including decapped deadenylated intermediates and NMD targets (Nagarajan et al., 2019).

In yeast, Xrn1 is involved in the degradation of the debranched intronic lariat (Hilleren and Parker, 2003). XRN1 is implicated in micro RNA degradation in humans (Bail et al., 2010). In a *xrn4* mutant in plants, siRNA targeting endogenous genes accumulate, resulting in gene silencing (Gazzani et al., 2004), suggesting that XRN4 also target RNAs that are substrate for RDR proteins (Tsuzuki et al., 2017). A similar role has been suggested for its homolog in *C. elegans* (Newbury and Woollard, 2004).

XRN proteins in yeast and Arabidopsis are involved in co-translational mRNA decay (Merret et al., 2015; Pelechano et al., 2015; Yu et al., 2016). In fact, Arabidopsis XRN4 co-translational mRNA decay is crucial for plant development (Carpentier et al., 2020). Recent papers, not yet peer-reviewed, also revealed that XRN4 regulates the accumulation of some circadian transcripts (Careno et al., 2022; Jones et al., 2022).

c. The 3'-5' degradation pathway

The exosome core and its activity

The RNA exosome is a key 3'-5' exoribonucleolytic activity in eukaryotes. The RNA exosome has a core complex containing nine subunits, called Exo9, interacting with various factors depending on the cellular context. Those factors include RNA helicases, RNA binding proteins, poly(A) polymerases and scaffold proteins. Exo9 is composed by three heterodimers of RNase PH-like domain proteins (Rrp41-Rrp45, Mtr3-Rrp42, Rrp46-Rrp43) structured as a ring and three S1 and/or KH RNA-binding proteins (Rrp4, Rrp40, Csl4) arranged as a cap-like structure onto the ring (Liu et al., 2006).

In fungi and metazoans, Exo9 lacks catalytic activity (Dziembowski et al., 2007; Schneider et al., 2009). By contrast, Arabidopsis Exo9 has a polyribonucleotide phosphorylase activity. It uses inorganic phosphate to cleave phosphodiester bonds with a distributive activity. Moreover, it can also synthetize RNA from nucleosides diphosphate (Sikorska et al., 2017). Sikorska *et al.* show that mutating RRP41 catalytic site completely abolishes Exo9 activity. Although Exo9 activity is not essential to Arabidopsis survival, the amino acids implicated in RNA binding, phosphate recognition and magnesium coordination are highly conserved in the plant lineage, suggesting that the peculiar activities of Exo9 could be conserved (Lange and Gagliardi, 2022; Sikorska et al., 2017). Moreover, the phosphorolytic activity of Exo9 might not be limited to the plant lineage since *RRP41* genes with intact active site have been found in the genomes of two Amoebozoa, *Capsaspora owczarzaki* and human pathogen *Naegleria fowleri* (Lange and Gagliardi, 2022).

3'-5' exoribonucleases associated with Exo9

- DIS3/RRP44

Dis3 is the main RNase associated with RNA exosomes in eukaryotes. It belongs to the RNase II/RNB family of proteins (Arraiano et al., 2013). Dis3 has a PIN domain at its N-terminus and a central RNA binding (RNB) domain. Dis3 binds to Exo9 via the PIN domain and the RNB domain confers processive 3'-5' exoribonucleolytic activity (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009; Zinder and Lima, 2017). In the nucleus and in the cytoplasm, Dis3 can degrade single stranded RNA and it can also unwind and degrade structured RNAs, if they have 3' single-stranded extensions (Wasmuth and Lima, 2012). It is interesting to note that Dis3 harbors both endo- and exoribonucleolytic activities (Chlebowski et al., 2011; Dziembowski et al., 2007).

Of note, three homologs of Dis3 have been found in humans: hDIS3, DIS3L1/DIS3L and DIS3L2. DIS3 also has endo- and exoribonucleolytic activities and associates with nuclear Exo9 (Chlebowski et al., 2011; Tomecki et al., 2010). DIS3L1/DIS3L localizes within the cytosol and on the contrary to DIS3, it does not have endoribonucleolytic activity (Tomecki et al., 2010). DIS3L2 on the other hand acts independently from the exosome and will be discussed later on.

In Arabidopsis, Exo9 associates with Rrp44/Dis3 homolog RRP44, which is essential for the viability (Kumakura et al., 2013). RRP44/DIS3 localizes both in the nucleus and the cytoplasm. Like its yeast homolog, it has a N-terminal PIN domain that mediates the interaction with Exo9 and a C-terminal RNB domain harboring the exoribonucleolytic activity (Kumakura et al., 2013).

- RRP6

Rrp6 is an exoribonuclease of RNase D-type (Briggs et al., 1998) with a hydrolytic distributive activity (Januszyk et al., 2011). It preferably acts on single-stranded RNA substrates but its activity on structured RNA with 3'-single stranded extension has also been shown. Rrp6 binds to the exosome core in the nucleus via its C-terminal part (Wasmuth and Lima, 2016). In yeast and human cells, Rrp6 is mostly localized within the nucleus (Burkard and Butler, 2000; Tomecki et al., 2010). Associated with Rrp6, the nuclear exosome degrades cryptic unstable transcripts *i.e*., non-coding RNAs produced from intergenic and intragenic regions, and non-coding RNA precursors, including ribosomal RNAs (Callahan and Butler, 2008; Gudipati et al., 2012; Preti et al., 2013; Sloan et al., 2013; Tafforeau et al., 2013)

On the contrary to yeast and metazoans, there are three RRP6-like proteins in Arabidopsis named RRP6L1, RRP6L2 and RRP6L3 (Lange et al., 2008). RRP6L1/2 localize in the nucleus whereas RRP6L3 is strictly located in the cytosol (Lange et al., 2008). While RRP6L2 has been shown to be an exosome cofactor, RRP6L1 is not (Lange et al., 2008; Zhang et al., 2014) and RRP6L3 function is not studied yet. RRP6L2 mostly localizes within nucleoli where it is involved in the maturation and degradation of rRNAs (Kumakura et al., 2013; Lange et al., 2008; Sikorska et al., 2017; Sikorski et al., 2015). In addition, RRP6L2 is able to trim 18S rRNA precursors, independently from the exosome (Sikorski et al., 2015).

Exosome cofactors in the nucleus

The decay activity of the exosome is regulated by cofactors, like RNA helicases from the Mtr4/Ski2 family, in a cell compartment-dependent manner. The exosome cofactors can unwind the RNA substrate or serve as scaffold proteins (Weick and Lima, 2021). In the nucleus, the RNA helicase Mtr4 binds to Exo9 via two adapter proteins, Mmp6 and Rrp47 (Schmid and Jensen, 2019) and its absence is lethal for yeast and human cells. Mtr4 is involved in several protein complexes between the exosome and other cofactors. Together with a zinc-finger protein (Air1 or Air2) and a non-canonical poly(A) polymerase (Trf4 or Trf5) that adenylates exosome substrates (Schmid and Jensen, 2019), Mtr4 belongs to the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) complex that degrade many types of non-coding RNAs in the nucleolus of yeast and human cells (Delan-Forino et al., 2020; Sudo et al., 2016; Vaňáčová et al., 2005). The TRAMP complex has not been reported in plants to date. In Arabidopsis, there are two putative homologs of Trf4/5, MEE44 and TRL. MEE44 function remains uninvestigated but TRL adenylation activity has been reported (Sikorski et al., 2015).

In the nucleoplasm of human cells, MTR4 is also part of the nuclear exosome targeting (NEXT) complex with the zinc-finger protein ZCCHC8 and the RNA-binding protein RBM7. NEXT regulates the exosome-mediated decay of some PROMPTs, enhancer RNAs, lncRNAs and 3'-extended snoRNAs or snRNAs (Andersen et al., 2013; Lubas et al., 2011, 2015).

By contrast with yeast and humans, two MTR4 paralogs exist in Arabidopsis genome, MTR4 and HEN2, that function in nuclear RNA surveillance (Lange and Gagliardi, 2022). Arabidopsis NEXT complex is composed of HEN2, the two ZCCHC8 homologs ZCCHHC8A/B and RBM7 (Bajczyk et al., 2020). To date, the substrates of plant NEXT complex remain unidentified, except for miRNA precursors and maturation by-products (Bajczyk et al., 2020). The NEXT complex can also act on newly transcribed RNAs with the assistance of the zinc-finger protein SERRATE and the nuclear cap-binding complex (Bajczyk et al., 2020; Lange et al., 2014). Plants mutated in SERRATE accumulate targets of HEN2, for

example precursors of miRNAs and snoRNAs, snRNAs, lncRNAs and short transcripts from protein-coding genes (Bajczyk et al., 2020). Hence, it has been proposed that SERRATE and the cap-binding complex are connected to the nuclear exosome via NEXT (Lange and Gagliardi, 2022).

Arabidopsis Exo9 also cooperates with MTR4 in the nucleolus (Lange et al., 2011) but how the association occurs remains unknown yet. This helicase copurifies with ribosomal biogenesis factors and rRNA precursors accumulate in *mtr4* mutant (Lange et al., 2011, 2014), so it has been proposed that MTR4 helps the exosome in ribosomal RNA processing and decay of maturation by-products (Lange and Gagliardi, 2022).

In human cells nuclei, the exosome is involved in the degradation of lncRNA, spliced transcripts from genes containing snoRNAs and truncated transcripts produced from protein-coding genes with internal poly(A) sites (Meola et al., 2016; Ogami et al., 2017). MTR4 interacts with PABPN1 via ZFC3H1 in the poly(A) exosome targeting (PAXT) connection (Meola et al., 2016). SOP1 has been proposed as putative ZCF3H1 homolog in Arabidopsis (Hématy et al., 2016). It colocalizes with HEN2 in nucleoplasmic speckles (Hématy et al., 2016). In this study, the authors observed an accumulation of non-spliced mRNAs, lncRNAs and primary miRNAs, that are HEN2 targets. HEN2 and SOP1 were proposed to be mobilized at sites of RNA 3'-end processing, potentially in a PAXT-like context (Lange and Gagliardi, 2022).

Exosome cofactors in the cytosol

In the cytosol, the Ski complex regulates exosome activity. In yeast, Ski7 connects Exo9 to the Ski complex that contains the RNA helicase Ski2, the tetratricopeptide repeat-containing protein Ski3 and two copies of the WD40-repeat protein Ski8 (Brown et al., 2000; Halbach et al., 2013; Kowalinski et al., 2016). In yeast and metazoans, the Ski complex helps to recruit the exosome to mRNAs (Araki et al., 2001; Orban and Izaurralde, 2005), in particular in the case of stalled ribosomes (Halbach et al., 2013; Kowalinski et al., 2016; Schmidt et al., 2016), but also as part of the nonsense-mediated or no-stop decay pathways (Arribere and Fire, 2018; Guydosh et al., 2017; Van Hoof et al., 2002; Mitchell and Tollervey, 2003). Recently, a new factor associated with a subpopulation of the Ski complex was identified in yeast, Ska1, specifically required for the degradation of long-3'UTR-containing RNAs, independently of a ribosome association (Zhang et al., 2019a).

In Arabidopsis, the cytosolic exosome also associates with the SKI complex via SKI7. The SKI complex consists of the RNA helicase SKI2, the tetratricopeptide repeat protein SKI3 and a dimer of the WD40 repeat protein SKI8/VIP3, as in other eukaryotes (Dorcey et al., 2012; Zhang et al., 2015b). As its metazoan homolog, SKI7 copurifies with exosomes and SKI complexes (Kalisiak et al., 2017; Lange et al., 2019; Zhang et al., 2015b), suggesting that SKI7 also connects the SKI complex to the core exosome in the cytosol.

Several studies in Arabidopsis have reported that fragments produced from RISC-cleaved transcripts are degraded by the cytoplasmic exosome, via the recruitment of the SKI complex (Branscheid et al., 2015; Vigh et al., 2022). A recent study has described the role of the plant exosome, along with the SKI complex and PELOTA proteins, in the degradation of non-stop decay substrates (Vigh et al., 2022). In addition, RST1 and RST1 INTERACTING PROTEIN (RIPR) also copurify with the Arabidopsis RNA exosome. These proteins are exclusively located within the cytosol and connect Exo9

and the SKI complex (Lange et al., 2019). Of note, RST1 homolog in mammals, FOCAD, also associates with the Ski complex, as well as the RNA binding protein AVEN (Tuck et al., 2020). The authors of this study proposed that AVEN prevents ribosome stalling and Ski recruitment to RNA structured, thereby hindering mRNA decay. In Arabidopsis, the molecular functions of RST1 and RIPR are not elucidated yet.

The 3'-5' exoribonuclease DIS3L2/SOV

Alternatively to the RNA exosome-mediated degradation, RNA can be degraded by a DIS3 homolog, the Dis3-like 3'-5' exoribonuclease 2 (DIS3L2). This protein from the RNase II/RNB family has been characterized in fission yeast and in metazoans (Astuti et al., 2012; Chang et al., 2013; Lin et al., 2017; Lubas et al., 2013; Malecki et al., 2013; Reimão-Pinto et al., 2016; Towler et al., 2016; Ustianenko et al., 2013; Weaver et al., 2014; Zhou et al., 2017). DIS3L2 contains two cold shock domains (CSD) at its N-terminus, a catalytic RNB domain and a S1 domain at its C-terminus (Lubas et al., 2013). The lack of PIN domain as compared to DIS3 could explain its exosome-independent activity. The RNB domain confers the exoribonuclease activity whereas the CSD and S1 domains are proposed to be RNA binding domains. *In vitro*, DIS3L2 is a processive enzyme and can degrade both single stranded and double stranded RNA, whether it is structured or not (Lubas et al., 2013; Ustianenko et al., 2013).

DIS3L2 localizes in the cytoplasm (Astuti et al., 2012; Lubas et al., 2013; Ustianenko et al., 2013) and targets many types of RNAs. DIS3L2-mediated degradation is described for miRNA and their precursors (Chang et al., 2013; Haas et al., 2016; Nowak et al., 2017; Ustianenko et al., 2013), as well as for other non-coding RNAs like long non-coding RNAs (Łabno et al., 2016; Pirouz et al., 2016; Ustianenko et al., 2016) and ribosomal RNAs (Pirouz et al., 2019).

Moreover, in human cells and fission yeast, mRNA can be degraded by DIS3L2 (Lubas et al., 2013; Malecki et al., 2013). NMD targets are also DIS3L2 substrates (da Costa et al., 2019; Kurosaki et al., 2018).

The mechanisms of action of DIS3L2 are intensively studied in fission yeast and human and mutations of *DIS3L2* have been reported to be associated with diseases. Mutations in *DIS3L2* have been found in patients with Perlman syndrome (Astuti et al., 2012; Morris et al., 2013), who are more prone to develop Wilms tumor (Astuti et al., 2012; Higashimoto et al., 2013; Soma et al., 2017; Wegert et al., 2015). In addition, it has been described that a knockout of *DIS3L2* stimulates cell proliferation and cancer cell growth (Astuti et al., 2012; Towler et al., 2016).

The existence of a DIS3L2 protein was originally identified in Arabidopsis and named SUPPRESSOR OF VARICOSE (SOV) (Zhang et al., 2010). In this study, the authors observed that plants from the Columbia (Col-0) accession mutated in *VARICOSE* (*VCS*), coding for a scaffold protein of the decapping complex (Xu et al., 2006), exhibit severe phenotypic defects, while mutant plants from the Landsberg (Ler) accession display mild phenotypes. A genetic screen identified a mutation in SOV locus of Col-0 as compared to Ler. So, by contrast to its homolog DIS3L2, the absence of SOV in Arabidopsis is not as deleterious to the plant as it is in animals (Astuti et al., 2012; Morris et al., 2013). Zhang and colleagues identified a point mutation in Col-0, resulting in the incorporation of an arginine at position 705 instead of a proline in SOV^{Ler} (Zhang et al., 2010). SOV participates in 3'-5' decay of

mRNAs and its activity is redundant with the exosome and the 5'-3' decay pathway (Sorenson et al., 2018).

d. Interplay between RNA degradation pathways and transcript buffering

5'-3' and 3'-5' pathways are not mutually exclusive. A recent study in mouse embryonic stem cells shows that both XRN1 and the Ski complex bind to translating ribosomes. Whereas XRN1 degrades full-length mRNAs as part of bulk mRNA turnover, the Ski complex functions in translation surveillance (Tuck et al., 2020).

In Arabidopsis, a transcriptome-wide study has shown that the 5'-3' pathway contributes to the decay of 68% of the transcriptome, in particular for protein-coding RNAs with short half-lives. The decay of about 75% of the transcriptome requires decapping complex scaffold protein VCS and/or 3'-5' exoribonuclease SOV. Only a few transcripts are direct substrates of SOV independently from decapping (Sorenson et al., 2018). Interestingly, some transcripts display faster decay rates in a *sov* background, without any changes in RNA abundance as compared to wild-type. The authors propose that the absence of SOV results in a feedback reaction to maintain RNAs at wild-type levels (Sorenson et al., 2018), comparable at a process that has already been observed in yeast: transcript buffering. Transcript buffering involves the shuttling of Xrn1 between the cytoplasm and the nucleus (Haimovich et al., 2013; Sun et al., 2013). As plant XRN4 is not reported to have a nuclear localization, the authors propose that other factors of mRNA decay are involved in transcript buffering in the *sov* background. Moreover, mutations in SOV are not lethal (Zhang et al., 2010), which could be explained by transcript buffering.

RNA degradation is a tightly regulated process that involves many different factors, some being highly conserved among eukaryotes. The regulation of gene expression is important for plant development, but is also an essential part of the defense against pathogens, such as viruses.

2. Interplay between RNA degradation factors and plant viruses

RNA decay mechanisms can also function in the antiviral response by degrading viral RNAs. In the next part, I will summarize our knowledge on the interactions between cellular RNA degradation factors and plant RNA viruses.

a. LSM1-7/PAT1 and decapping complex factors are needed for viral replication

In plants, the brome mosaic virus (BMV, *Bromoviridae*) interplay with RNA degradation factors has been broadly investigated. Some of these proteins are essential for BMV infection such as the LSM1-7 complex, PAT1 and the RNA helicase RH12 (Alves-Rodrigues et al., 2007; Díez et al., 2000; Galão et al., 2010; Jungfleisch et al., 2015; Kushner et al., 2003; Mas et al., 2006; Noueiry et al., 2003). Mutations of these factors result in a decrease in viral RNA replication and translation. As for mRNA, the LSM1-7/PAT1 complex binds to the 3' extremity of the viral RNA but it does not need an oligo(A), as it has been shown that they directly bind to sequences within the 3'UTR of BMV RNAs (Galão et al., 2010; Jungfleisch et al., 2016; Kushner et al., 2003). In association with translation initiation factors interacting with BMV RNA 5'UTR, these factors help in viral translation and replication. Interestingly, if the 3'UTR

of BMV RNAs is replaced by a poly(A) tail, the viral translation and replication no longer needs the LSM1-7 complex (Mas et al., 2006; Noueiry et al., 2003).

A recent study on TuMV has demonstrated that LSM1 and PAT proteins are also involved in TuMV infection. Both are upregulated upon TuMV infection and their function in mRNA decay seems to be altered, as the authors show an accumulation of two mRNA decapping targets. In addition, they observed less viral capsid protein in plants mutated for these factors. As TuMV RNA level does not decrease, they propose that LSM1 and PAT1 are hijacked by TuMV in its own interest (Zuo et al., 2022).

A decapping complex component, VCS, has been reported to be needed for potato virus A (*Potyviridae*) replication (Hafrén et al., 2015)

b. Complex roles of XRN4 in viral infection

Antiviral role of XRN4

Silencing of XRN4 results in a faster and broader systemic infection of TMV in *N. benthamiana* (Peng et al., 2011) and promotes the tombusvirus TBSV RNA accumulation (Jaag and Nagy, 2009). One study observed an increased level of the potexvirus PlMAV and the potyvirus TuMV in *xrn4-5* mutant (Ma et al., 2019). Overexpression of OsXRN4 in *Oryza sativa* (rice) and *N. benthamiana* leads to reduced level of viral RNA of rice stripe virus and TMV, a ss(-) RNA and ss(+) RNA virus, respectively (Jiang et al., 2018). XRN4 overexpression also enhances the viral RNA degradation of the tombusvirus cucumber necrosis virus (Cheng et al., 2007). In a study focusing on potexviruses, a slight decrease in the accumulation of the foxtail mosaic virus has been observed upon XRN4 overexpression in *N. benthamiana*, and no effect has been seen on the accumulation of PVX (Lee et al., 2015). By contrast, another study observed that PVX accumulation is reduced upon XRN4 overexpression in *N. benthamiana*, as well as the accumulation TuMV and PlAMV, two potyviruses (Ma et al., 2019). TuMV RNA accumulation has also been reported to be inhibited by co-expression of XRN4 in *N. benthamiana*. The authors of this study also have shown that TuMV HC-Pro can directly interact with XRN4, so they suggest that this interaction limits XRN4 degradation activity (Li and Wang, 2018). Overall, these results indicate that XRN4 could be involved in viral RNA degradation for several plant RNA viruses.

Viral RNA features to resist 5'-3' degradation

Some viral RNAs display specific structures that blocks the action of XRN proteins, as it has been reported for several flaviviruses (Chapman et al., 2014; Funk et al., 2010; Moon et al., 2012; Pijlman et al., 2008; Silva et al., 2010).

These structures are not limited to animal viruses. A nuclease-resistant motif, named "coremin" motif, has been described in the sgRNA5 of a particular CMV strain (Thompson et al., 2008). A similar motif present in BNYVV RNA3 is resistant to nuclease attacks and essential for long distance movement (Peltier et al., 2012). This motif stalls XRN proteins and allows the production of a non-coding RNA (ncRNA3) that strongly accumulates. This may be a way to sequestrate XRN 5'-3' exoribonucleases. Interestingly, the down-regulation of XRN proteins in *N. benthamiana* has been reported to hinder BNYVV systemic movement, suggesting that XRN proteins are required for cell-to-cell movement by participating in the generation of the ncRNA3 (Flobinus et al., 2018). This "coremin" motif is actually

conserved between two benyviruses, the BNYVV and the beet soil-borne mosaic virus (Ratti et al., 2009).

Such structures have been frequently discovered in *Tombusviridae* (Gunawardene et al., 2019, 2021; Ilyas et al., 2021; Iwakawa et al., 2008; Steckelberg et al., 2018b, 2018a) but also in *Sobemoviridae* (Steckelberg et al., 2018a). These structures inhibit XRN-mediated degradation without the help of other factors (Steckelberg et al., 2018b).

Moreover, a recent study highlights that this motif is widespread in viral RNAs from *Betaflexviridae, Potyviridae, Secoviridae and Virgaviridae* species (Dilweg et al., 2019). The authors showed for representants from the latter viral families that this motif can resist XRN-mediated degradation *in vitro* and that it folds into two stem-loop structures. Of note, the structure described for the RNA of the carnation Italian ringspot virus (*Tombusviridae*) folds into a structure that prevents the binding of XRN4 by protecting the 5' phosphate, rather than stalling the exoribonuclease. The authors called it XRN-evading RNA (Gunawardene et al., 2021).

Proviral role of XRN4

Other studies reported a lower viral accumulation in the absence of XRN4. A genetic screen has identified a mutation in Arabidopsis XRN4 that leads to a decreased susceptibility to TuMV infection (Vogel et al., 2011). In a *xrn4-6* mutant, the cucumber mosaic virus (*Bromoviridae*) replication was shown to be reduced, as well as viral siRNAs level (Gy et al., 2007). Reduced accumulation of viral genomic and subgenomic RNAs upon silencing of XRN4 have been reported for the bamboo mosaic virus (*Potexvirus, Alphaflexviridae*) (Lee et al., 2015).

Several studies investigated the role of XRN 5'-3' exoribonucleases during viral infection. Most of them rely on the silencing or overexpression of XRN4 in *N. benthamiana* and viral infection has been quantified by detecting the capsid in a few plants by western blot of the capsid or by RT-PCR on the viral RNA. Some of these studies concluded that XRN4 is an antiviral factor because the authors observed a reduced or increased viral infection upon XRN4 overexpression or mutation, respectively. Yet, other studies have reported that XRN4 could have a proviral function because the virus accumulates less in a *xrn4* mutant, suggesting XRN4 could function in the viral metabolism.

In a *xrn4* mutant, siRNAs directed against endogenous transcripts accumulate (Gazzani et al., 2004; Gy et al., 2007), so the broader/lower viral infection could result from the silencing of host genes involved in the antiviral response or required for the viral infection, rather than from the direct viral RNA stabilization/destabilization. Moreover, these studies quantified the viral accumulation by western blot on the capsid protein or northern blot on viral RNAs often in a few plants. Some of their conclusions are drawn from visual observations, so they may lack reproducibility. These data suggest that further work will be required to conclude on XRN4 impact on plant viral RNA accumulation.

c. Other RNA degradation factors and their impact on viral accumulation

The expression of decapping factors DCP1 and DCP2 and PARN deadenylase in *N. benthamiana* has been reported to be upregulated upon TuMV infection (Li and Wang, 2018). The transient expression of these factors in *N. benthamiana* inhibits TuMV RNA accumulation. In line with this observation, knocking down one of these proteins results in an enhanced viral infection (Li and Wang,

2018). One weakness in this study is that they quantified viral accumulation from infiltrated leaves with proteins transiently expressed for a few plants. Another study has reported that there is no difference in viral accumulation in DCP2 mutant hypomorphic allele upon TuMV (*Potyvirus*) and PlAMV (*Potexvirus*) infections (Ma et al., 2019). The impact of decapping factors on plant viral accumulation remains to be investigated more closely.

Using transgenic plants expressing P1/HC-Pro from TuMV, one study identified SKI8/VIP3, component of the SKI complex, as an interactant of P1. The expression of P1/HC-Pro results in a strong serrated leaves phenotype, characteristic of a PTGS factor mutation. The authors thus propose that P1 may stimulate HC-Pro mediated PTGS suppression (Hu et al., 2020). P1, via its binding to SKI8 may interfere with SKI8 function in the degradation of 5' fragments generated from RISC-cleaved transcripts (Branscheid et al., 2015; Orban and Izaurralde, 2005).

An antiviral function of NMD has been proposed in plants (Garcia et al., 2014). Mutations in the main NMD factor UPF1 upon PVX infection results in higher viral protein and RNA levels, in particular for sgRNA3. This viral RNA has a long 3'UTR and is more sensitive to NMD than the other sgRNA produced during PVX infection, which does not possess a long 3'UTR. Moreover, NMD also hinders TCV infection but not TuMV infection. PVX and TCV both produces sgRNAs during their replication cycle, but TuMV does not. Instead, a polyprotein is translated and the viral proteins are generated by proteolytic cleavage. The polyprotein could prevent long 3'UTRs and internal termination codons, thereby avoid the recognition by NMD factors. An additional study revealed that TCV presented genomic features that confers some degree of resistance to NMD. TCV and other carmoviruses possess an unstructured region in their 3'UTR that is able to protect a reporter gene from NMD (May et al., 2018). Interestingly, similar unstructured regions are found widespread among RNA viruses and in NMD-resistant transcripts with long 3'UTRs in human, suggesting that they represent a common strategy to avoid NMD.

5. RNA uridylation

In eukaryotes, RNA 3' termini can be modified by the addition of untemplated nucleotides. The most frequent 3' untemplated nucleotide addition is polyadenylation. Uridylation is the second major 3' terminal modification detected in eukaryotes. It targets both coding and non-coding RNAs in all the eukaryotes tested, except *S. cerevisiae* (De Almeida et al., 2018b; Liudkovska and Dziembowski, 2021; Warkocki et al., 2018a; Yu and Kim, 2020; Zigáčková and Vaňáčová, 2018).

1. RNA uridylation: a key player of post-transcriptional regulation

a. Structural organization of the terminal uridylyl transferases

RNA 3' uridylation is catalyzed by terminal uridylyl transferases (TUTases) that belong to the DNA polymerase beta-like nucleotidyltransferase superfamily (De Almeida et al., 2018a).

The minimal configuration of a TUTase is a nucleotidyltransferase domain (NTD) and a poly(A) polymerase-associated (PAP-assoc) domain that forms the catalytic domain, as it has been described

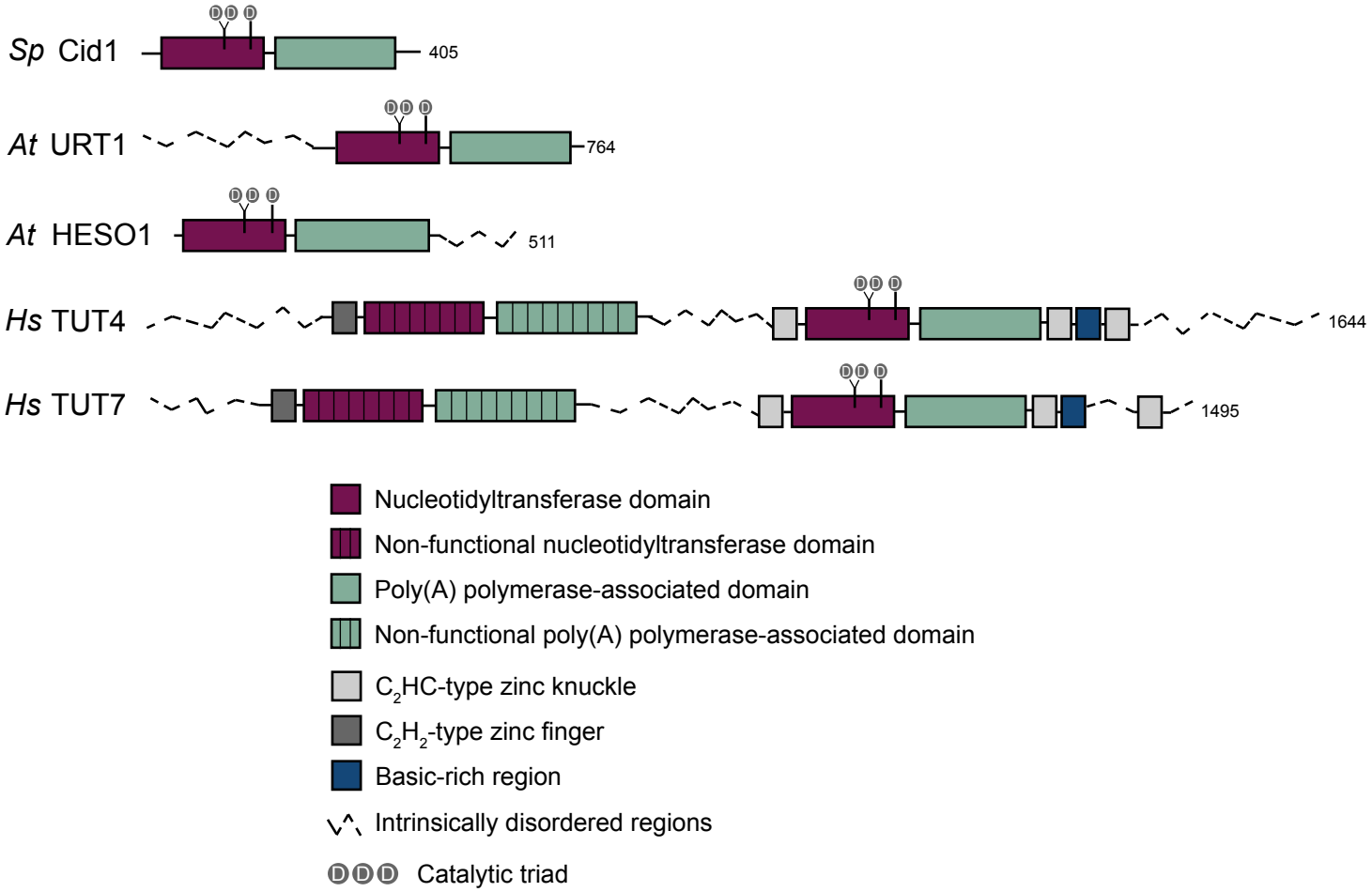


Figure 10: Domain organization of the TUTases is diverse. The structural organization is presented for five TUTases from *S. pombe* (*Sp*), *Arabidopsis thaliana* (*At*) and *Homo sapiens* (*Hs*). The three aspartic acid residues of the catalytic site are indicated by the code letter D. The number of residues of each TUTase is written on the right. Adapted from De Almeida *et al.*, 2018.

for TUTase Cid1 in fission yeast *S. pombe* (Fig. 10) (Munoz-Tello et al., 2012). Resolution of TUTase atomic structures revealed that the NTD and the PAP-assoc domain constitute a large cleft which contains the uridine 5'-triphosphate (UTP) binding and the catalytic domains (Munoz-Tello et al., 2012; Rajappa-Titu et al., 2016; Stagno et al., 2010; Yamashita et al., 2017; Yates et al., 2012, 2015; Zhu et al., 2020). It appears that the PAP-assoc domain has evolved to bind UTP instead of ATP. This specificity is conferred by a single glutamate or aspartate residue in Cid1 and *T. brucei* TUTase RET1, respectively (Stagno et al., 2007b, 2007a; Yates et al., 2015). Cid1's UTP specificity is also granted by the histidine at position 336, conserved in some plant and human TUTases (Yates et al., 2015). In Arabidopsis, there are two TUTases described so far, HEN1 SUPPRESSOR 1 (HESO1) (Ren et al., 2012) and UTP:RNA URIDYLYL TRANSFERASE 1 (URT1) (Sement et al., 2013). It is interesting to note that the histidine residue reported to be crucial for UTP selection is conserved in URT1 but not in the HESO1 sequence that contains a valine residue instead (Zhu et al., 2020). Hence, the nucleotide specificity could be less strict for HESO1 as for URT1. HESO1 was indeed reported to have a cytidylation activity (Song et al., 2019).

Except for SpCid1, all TUTases characterized to date have additional domains to the catalytic domain (Fig. 10). Human TUTases TUT4 and TUT7 have their catalytic domain duplicated, but only the C-terminal one remains active (Blahna et al., 2011; Faehnle et al., 2017). They also contain zinc finger motifs as well as intrinsically disordered regions (IDRs) (De Almeida et al., 2018b). In fact, almost all the TUTases described so far possess IDRs, that could participate in protein-protein interactions or RNA substrate recognition. For example, a motif in URT1 IDR has been shown to be essential for the interaction with the translation inhibitor/decapping activator DCP5 and it connects URT1 to additional decay factors (Scheer et al., 2021).

b. Multifaceted role of non-coding RNA uridylation

RNA uridylation emerged as a central player of the post-transcriptional regulation in eukaryotes and it has many functions in RNA maturation and degradation. I will describe the main functions of RNA uridylation described in eukaryotes, from non-coding RNA maturation and decay to mRNA degradation. Table 2 recapitulates the main TUTases identified so far and their functions.

Organism	TUTase	Function
T. brucei	RET1	guide RNA maturation in mitochondria
S. pombe	Cid1	mRNA degradation
A. nidulans	CutA	mRNA degradation
C. elegans	CDE-1	siRNA and viral RNA degradation, ribosomal RNA surveillance
D. melanogaster	Tailor	siRNA and mRNA degradation, non-coding RNA surveillance
H. sapiens	TUT1	U6 snRNA maturation
	TUT4	mRNA, LINE-1 retrotransposon and viral RNA degradation, let-7 precursor maturation or degradation, non-coding RNA surveillance
	TUT7	mRNA, LINE-1 retrotransposon and viral RNA degradation, let-7 precursor maturation or degradation, non-coding RNA surveillance
A. thaliana	URT1	mRNA degradation, miRNA and RISC-cleaved transcript degradation
	HESO1	siRNA, miRNA and RISC-cleaved transcript degradation
C. reinhardtii	MUT68	small RNA degradation

Table 2: Examples of TUTases identified in eukaryotes and their functions.

Maturation non-coding RNAs

In metazoans, there are two well studied examples of non-coding RNA maturation involving uridylation. The first example is the dual function of uridylation in let-7 miRNA biogenesis. Members of this miRNA family are involved in cell proliferation suppression and cell differentiation stimulation in animals (Su et al., 2012). Let-7 precursors either have a 2 nt 3' overhang (group I) or a 1 nt 3' overhang (group II). The latter are not well suited to be processed by Dicer and their processing relies on TUTase activity (Ha and Kim, 2014). In differentiated cells, the redundant activities of TUT4 and TUT7 (Thornton et al., 2012) mono-uridylate group II pre-let-7, thus reestablishing the processing activity by Dicer (Heo et al., 2012; Kim et al., 2015). Uridylation of let-7 precursor is also associated with its degradation. In embryonic stem cells and some cancer cells, the expression of let-7 is negatively regulated by Lin28 proteins (Balzeau et al., 2017). Lin28A binds to pre-let-7 miRNA and together with the E3 ligase Trim25, they recruit TUT4/7 that oligo-uridylate let-7 precursors (Hagan et al., 2009; Heo et al., 2012; Thornton et al., 2012; Wang et al., 2017). The oligo-uridylated let-7 precursors are subsequently degraded by DIS3L2 (Chang et al., 2013; Faehnle et al., 2014; Ustianenko et al., 2013). The example of let-7 helps to understand the complex role of RNA uridylation: mono-uridylation stimulates Dicer processing, thus downregulating the expression of genes involved in cell proliferation, whereas pre-let-7 oligouridylation induces its degradation. Let-7 has been found to be repressed in several cancers in mammals (Balzeau et al., 2017).

The second-best example of uridylation involved in non-coding RNA maturation in metazoans is the maturation of spliceosomal U6 small nuclear RNA (U6 snRNA) in humans. U6 snRNA is transcribed by RNA Polymerase III and the transcript ends with four encoded uridines, that are bound by the La protein. This allows the recruitment of TUT1 that uridylates U6 snRNA 3' extremity (Trippe et al., 1998, 2003, 2006). The oligo(U) is next nibbled by the 3'-5' exoribonuclease Ubs1 that generates a 3' terminal nucleotide with 2',3'-cyclic phosphate (Hilcenko et al., 2013; Mroczek et al., 2012; Shchepachev et al., 2012). The particular 3' extremity stimulates the binding of the LSm2-8 complex (Licht et al., 2008). For

U6 snRNA, terminal uridylation allows to complete its processing. In addition, the bound LSm2-8 complex prevents 3' adenylation by Trf4 and subsequent degradation by the nuclear exosome (Hilcenko et al., 2013).

Another interesting case of RNA maturation involving uridylation takes place in the parasitic kinetoplastid *Trypanosoma brucei*. In the mitochondria, the genome is widely edited by uridine-insertion/deletion and these editions are mediated by guide RNAs (gRNAs) (Aphasizheva and Aphasizhev, 2016). gRNAs are processed by the MPsome complex, constituted of the TUTase RET1, the 3'-5' exoribonuclease DSS1 and three large proteins (Aphasizhev et al., 2016; Suematsu et al., 2016). The gRNA forms a duplex between a sense and an antisense strand that has 3' overhang extremities. These are uridylated by RET1, inducing the binding of DSS1 that trims until 10-12 nt away from the duplex. Another round of uridylation is performed by RET1, promoting MPsome disengagement from the gRNA duplex intermediate. After unwinding the duplex, the antisense strand is degraded whereas the sense strand is integrated into the gRNA binding complex, which mechanism is not fully understood to date (Aphasizhev et al., 2016).

Uridylation-mediated non-coding RNA decay

Besides its involvement in RNA maturation, uridylation is a degradation signal on many types of small RNAs from animals to plants. Plants miRNAs and siRNAs are 2'O-methylated on their terminal ribose by HUA ENHANCER 1 (HEN1) (Li et al., 2005; Yang et al., 2006b; Yu et al., 2005). This modification stabilizes small RNAs, including miRNAs because the lack of 3' methylation causes a decrease in siRNA and miRNA abundance, resulting in strong developmental defects. In the absence of HEN1, small RNAs are uridylated and trimmed at their 3' extremity (Li et al., 2005; Yu et al., 2005). So, 3' methylation inhibits uridylation and subsequent decay of small RNAs in plants. The phenotype observed in *hen1* mutant is suppressed by a mutation at locus coding for a TUTase, called HEN1 SUPPRESSOR 1 (HESO1) and the phenotypic rescue does not rely on 3' methylation (Ren et al., 2012). In this work, the authors have shown that unmethylated miRNAs and siRNAs are accessible to HESO1. In the absence of HESO1 and HEN1, the global level of small RNAs increases so the authors proposed that terminal uridylation destabilizes small RNAs in plants (Ren et al., 2012; Zhao et al., 2012). Of note, the second TUTase identified in Arabidopsis, URT1 can uridylate certain miRNAs, but mostly in the absence of HEN1 and HESO1 (Wang et al., 2015). Terminal methylation prevents HESO1 activity and uridylation has been reported on trimmed small RNAs, hence small RNAs are trimmed prior to uridylation (Ren et al., 2012; Zhai et al., 2013; Zhao et al., 2012). Methylated small RNAs are trimmed by the 3'-5' exoribonucleases SDN1 and SDN2 (Yu et al., 2017). HESO1 and SDN1/2 both interacts with the silencing complex via AGO1 (Ren et al., 2014; Yu et al., 2017). Of note, the TUTase MUT68 in *C. reinhardtii* also uridylates small RNAs (Ibrahim et al., 2010). Although it is reported that 3' uridylation destabilizes plant small RNAs, the degradation activity of uridylated small RNAs in plants remains unknown yet.

In metazoans, 3' methylation has been detected on Piwi-interacting RNAs (piRNAs) in *D. melanogaster, D. rerio, C. elegans, M. musculus*, human cells and the unicellular eukaryote Tetrahymena (Horwich et al., 2007; Kamminga et al., 2010; Kirino and Mourelatos, 2007; Kurth and Mochizuki, 2009; Lim et al., 2015; Montgomery et al., 2012; Saito et al., 2007). Methylation of piRNAs

has been proposed to prevent uridylation and subsequent decay (Kamminga et al., 2010; Kurth and Mochizuki, 2009; Lim et al., 2015). Terminal methylation has also been reported of drosophila and trypanosoma small interfering RNAs (siRNAs) (Horwich et al., 2007; Shi et al., 2014). In *D. melanogaster*, the lack of 3' methylation of siRNAs is described to induce their trimming and uridylation (Ameres et al., 2010). These results indicate that in eukaryotes, some small RNAs are methylated at their 3' end and it protects them from terminal uridylation and subsequent decay.

The decay of mature miRNAs in complex with AGO also involves uridylation in animals. They are uridylated by TUT4/7 and then degraded by DIS3L2 (Yang et al., 2020). In fact, uridylation has been detected on many types of small RNAs and their precursors and emerges as an important regulator of their accumulation in metazoans and *S. pombe* (Gutiérrez-Vázquez et al., 2017; Jones et al., 2009, 2012; Kim et al., 2015; Knouf et al., 2013; Newman et al., 2011; Pirouz et al., 2016; Pisacane and Halic, 2017; Thornton et al., 2014; van Wolfswinkel et al., 2009; Wyman et al., 2011). The consequence of uridylation is not investigated for all of them, but at least in the case of miRNA, uridylation reduces their abundance in human cells (Baccarini et al., 2011; Kim et al., 2015; Knouf et al., 2013; Thornton et al., 2014). Of note, 3' terminal uridylation enhances DIS3L2 recruitment but it is not necessary, as DIS3L2 can function in a polyuridylation-independent manner (Nowak et al., 2017).

Recently, a study revealed that the non-coding RNA 7SL, truncated and uridylated at its 3' terminus, accumulate in DIS3L2-depleted cells (Pirouz et al., 2020). 7SL RNA is part of the signal recognition particle complex that controls co-translational translocation of proteins to the ER (Walter and Blobel, 1982). The uridylated form of 7SL RNA that accumulates upon DIS3L2 depletion blocks the function of the signal recognition particle complex. Hence, DIS3L2-mediated decay has been proposed to regulate the ER-targeted mRNA translation and subsequent ER-targeted calcium homeostasis (Pirouz et al., 2020)

Moreover, uridylation is associated to non-coding RNA surveillance *i.e.*, the decay of misprocessed non-coding RNAs. It was reported in *D. melanogaster, M. musculus* and human cells that uridylated non-coding RNAs are substrates of the 3'-5' exoribonuclease DIS3L2 (Łabno et al., 2016; Pirouz et al., 2016; Reimão-Pinto et al., 2016; Ustianenko et al., 2013, 2016). In *D. melanogaster*, the TUTase Tailor and Dis3l2 associates into the Terminal RNA Uridylation-Mediated Processing (TRUMP) complex to participate in the degradation ribosomal RNAs, small nuclear RNAs and transfer RNAs (Reimão-Pinto et al., 2016). The unprocessed non-coding RNAs present in the cytosol of mammals are uridylated by TUT4/7, mostly at position close to stable secondary structures (Łabno et al., 2016; Pirouz et al., 2016; Ustianenko et al., 2016). A recent study in *C. elegans* also highlighted the role of uridylation, cooperating with DIS3L2 homolog SUSI-1, in ribosomal RNA surveillance (Wang et al., 2020).

c. The complex role of mRNA uridylation

The case of replication-dependent histone mRNAs in mammals

Uridylation is not limited to non-coding RNAs and has also been detected on mRNAs. In mammals, the maturation of histone protein-coding mRNAs involves uridylation. This particular 3' end processing is tightly regulated and essential for the production of high levels of histones only during the S-phase of the cell cycle (Marzluff and Koreski, 2017). On the contrary to the other mRNAs that are polyadenylated,

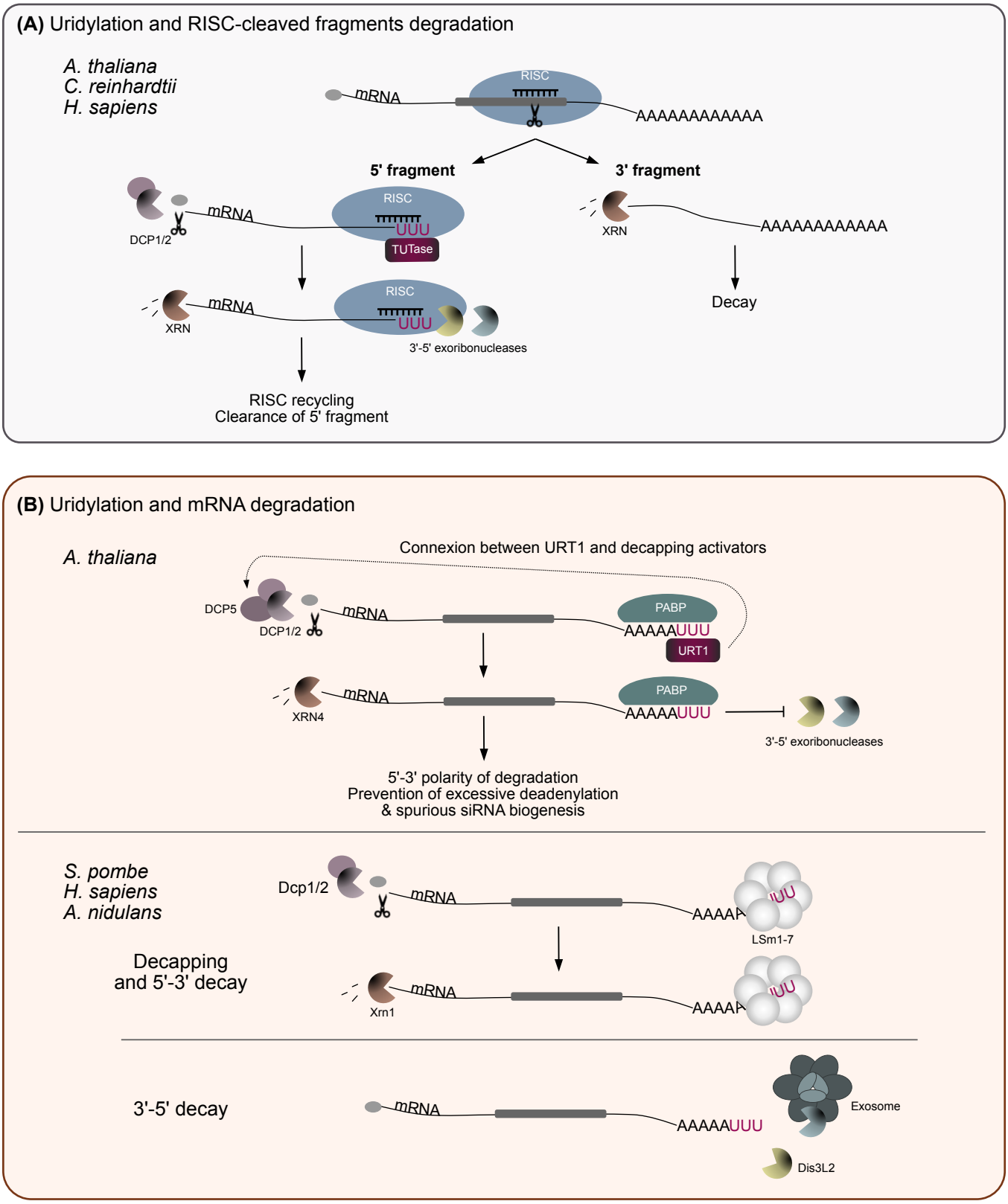


Figure 11: Uridylation and messenger RNA degradation. (A) RISC-cleavage generates two fragments. The 3' fragment is degraded by XRN proteins while the 5' fragment is uridylated. Uridylation facilitates RISC recycling and decay of the fragment. **(B)** After a deadenylation step, mRNAs are uridylated. In *A. thaliana*, uridylation restores PABP binding site and the main TUTase URT1 associates with decapping activators. Uridylation prevents excessive deadenylation and polarizes the degradation from 5'-3'. In other eukaryotes, uridylated mRNAs are bound by the LSm1-7 complex that triggers decapping and subsequent degradation by Xrn1. Uridylated mRNAs can alternatively be degraded from their 3' end by the exosome or Dis3L2.

the transcript 3' end folds into a stem loop structure (Marzluff and Koreski, 2017; Pandey and Marzluff, 1987). The stem loop is bound by the 3'-5' exoribonuclease hExo (Eri1) and the Stem-Loop Binding Protein (SLBP) (Tan et al., 2013). During the S-phase, additions of one or two uridines have been detected at histone mRNA terminal stem loops. These additions are proposed to restore the functional length after 3' hExo trimming and to stabilize histone mRNAs (Lackey et al., 2016; Welch et al., 2015). Histone mRNAs synthesis must be quickly stopped at the end of the S-phase and the arrest relies on uridylation. hExo removes 5-7 nt and the mRNA is oligouridylated by TUT7 (Hoefig et al., 2013). Successive rounds of trimming and uridylation occurs until hExo reaches the inside of the stem-loop, where it cannot bind to the RNA anymore. Upf1, key player of the non-sense mediated decay (Kaygun and Marzluff, 2005), has been proposed to interact with SLBP bound to the stem-loop and to stimulate 3'-5' decay by the exosome (Slevin et al., 2014). It appears that the 5'-3' decay pathway does not have a major impact on histone mRNA degradation, although it can occur (Mullen and Marzluff, 2008; Song and Kiledjian, 2007). As for let-7 precursors, uridylation plays a dual role on histone mRNAs, stabilizing them with the addition of 1-2 Us and triggering their decay with their oligouridylation.

mRNA fragment degradation

Uridylation plays an important role in the degradation of mRNA fragments in the cell. The 5' fragment resulting from RISC-mediated cleavage is also uridylated and has been detected in plants and murine cells (Shen and Goodman, 2004; Xu et al., 2016). In Arabidopsis, both TUTases can uridylate these fragments, with a predominant role of HESO1 (Ren et al., 2014; Zuber et al., 2018). Two 3'-5' exoribonucleases involved in the initiation of RISC-cleaved fragment decay have been characterized, RICE1 and RICE2. The authors show that they both interact with AGO1 and AGO10 and propose that RICE1/2 help RISC recycling (Zhang et al., 2017). Here, uridylation is a signal for degradation which is carried out by XRN4 (Shen and Goodman, 2004) or the RNA exosome (Branscheid et al., 2015; Vigh et al., 2022). Uridylation and subsequent degradation of RISC-cleaved 5' fragment is conserved in the green algae *C. reinhardtii* (Ibrahim et al., 2006). In metazoans, uridylation of RISC-cleaved fragments is ensured by TUT2 (Xu et al., 2016) and as in plants, uridylation triggers degradation either by XRN proteins or the RNA exosome (Orban and Izaurralde, 2005; Song and Kiledjian, 2007) (Fig. 11A).

Other mRNAs fragments generated during cell apoptosis are uridylated by TUT4/7 and further degraded by the 3'-5' exoribonuclease DIS3L2 (Thomas et al., 2015). Uridylation has recently been observed on several fragments originating from 5' UTRs, called 5' mRFs (Łabno et al., 2016; Ustianenko et al., 2016). As 5' mRFs with 3' terminal uridylation accumulate in cells overexpressing catalytically inactive DIS3L2, it suggests that these fragments normally undergo degradation by DIS3L2, stimulated by uridylation (Łabno et al., 2016; Ustianenko et al., 2016). The cooperation between TUTases and DIS3L2 helps to clear transcription by-products exported into the cytoplasm, but this pathway remains to be more investigated.

Recently, a study has also reported the impact of 3' uridylation on the regulation of LINE-1 retrotransposon. The authors have detected pervasive uridylation on LINE-1 mRNAs, resulting in the inhibition of LINE-1 retrotransposition. TUTases TUT4/TUT7 cooperate with the RNA helicase MOV10. Both TUTases have differential activities: TUT4-mediated uridylation destabilizes LINE-1 mRNAs

whereas the uridines added by TUT7 hinder the initiation of reverse transcription (Warkocki et al., 2018b).

Multiple aspects of uridylation in polyadenylated mRNA degradation

Uridylation on polyadenylated mRNAs has been reported in *S. pombe*, the fungus *A. nidulans*, *A. thaliana*, mammals, *T. brucei* mitochondria, and *X. laevis* and *A. pectinifera* oocytes (Aphasizheva et al., 2011; Chang et al., 2014; Lapointe and Wickens, 2013; Morgan et al., 2017; Morozov et al., 2010, 2012; Ochi and Chiba, 2016; Rissland and Norbury, 2009; Sement et al., 2013).

The initial function described for poly(A) mRNA uridylation is to induce decay in *S. pombe* via the recruitment of the LSm1-7 complex that stimulates the decapping and 5'-3' exoribonuclease activity (Rissland and Norbury, 2009) (Fig. 9). In *S. pombe* and mammalian cells, 3' terminal uridylated deadenylated transcripts are enriched in *lsm1* mutants (Lim et al., 2014; Rissland and Norbury, 2009), suggesting that 3' terminal uridylation favors the binding of the Lsm1-7/Pat1 complex. The role of uridylation in stimulating Lsm1-7 recruitment, and subsequent decapping, has been proposed in *S. pombe* (Fig. 9, 11) (Rissland and Norbury, 2009). In addition, it has been reported that 3' terminal uridylation stimulates the binding of the Lsm1-7, and subsequent decapping in mammalian cells (Song and Kiledjian, 2007).

3' terminal uridylation can favor the 3'-5' degradation by DIS3L2 in mammals (Lim et al., 2014; Malecki et al., 2013). By contrast to *S. pombe*, uridylation is preceded by deadenylation in other eukaryotes as many studies report uridylation on short poly(A) tails at a transcriptome-wide or single-mRNA level (Chang et al., 2014; Morozov et al., 2010, 2012; Scheer et al., 2021; Sement et al., 2013; Zuber et al., 2016). It is interesting to note that mRNA uridylation is widespread in human cells (Chang et al., 2014) and in Arabidopsis (Scheer et al., 2021; Zuber et al., 2016) as a percentage of all mRNAs are uridylated.

In human cells and *A. nidulans*, mRNA uridylation can also triggers degradation by the 5'-3' exoribonuclease XRN1 or by facilitating the recruitment of nonsense-mediated decay factors (Morozov et al., 2010, 2012; Sharif and Conti, 2013; Song and Kiledjian, 2007). In Arabidopsis, decapping can also be stimulated by uridylation via the TUTase URT1 that interacts with decapping activator DCP5, ortholog to human LSM14 (Fig. 9, 11) (Scheer et al., 2021).

Developmental processes are tightly regulated by uridylation, as reported by studies in mammals and *C. elegans*. In mammals, uridylation is a crucial to shape both maternal and male germline transcriptomes. Two studies reported that TUT4/7 are responsible for 3' terminal uridylation of mRNAs during oocyte growth to eliminate transcripts (Morgan et al., 2017) as well as during meiotic progression in spermatocytes (Morgan et al., 2019). In *C. elegans*, poly(U) polymerases (PUPs) are essential for germline development (Li and Maine, 2018).

In Arabidopsis, our team largely contributed to the investigation of the mechanisms related to RNA uridylation, in particular on mRNAs. URT1 has been identified as the main TUTase responsible for uridylating mRNAs (Sement et al., 2013). In addition, uridylation prevents the accumulation of excessively deadenylated mRNAs (Scheer et al., 2021; Sement et al., 2013; Zuber et al., 2016). At a transcriptome-wide scale, URT1 participates in shaping the poly(A) tail sizes in Arabidopsis (Scheer et al., 2021). In addition, uridylation of deadenylated mRNAs restores an average length of 16 nt, which is

enough for the binding of PABP (Zuber et al., 2016). Interestingly, uridines at the 3' end of the poly(A) tail hinder CAF1b deadenylation activity *in vitro* (Scheer et al., 2021). Additionally, upon transient overexpression of URT1, the length of the poly(A) tail increases (Scheer et al., 2021; Zuber et al., 2016). These results indicate that uridylation and deadenylation/polyadenylation are interconnected. Moreover, URT1 copurifies with decapping activators and translation inhibitors such as VCS or DCP5 (Scheer et al., 2021). Altogether, these results suggest that mRNA uridylation in Arabidopsis could participate in the 5'-3' polarity of degradation, by stimulating decapping and preventing excessive deadenylation (Fig. 11B). Although the plant does not exhibit phenotypic defects in the absence of URT1, the simultaneous mutation of URT1 and XRN4 has a strong negative impact on plant fitness. Our team showed that this phenotype was induced by an accumulation of spurious siRNAs against endogenous mRNAs. So, they propose that, by preventing the accumulation of excessively deadenylated mRNAs, URT1-mediated uridylation hinders the production of spurious siRNAs targeting mRNAs (Scheer et al., 2021).

Uridylation plays a central role in the post-transcriptional regulation of coding and non-coding RNAs in eukaryotes. It is involved in the maturation and degradation of several RNA species, that take part in many molecular processes.

2. Viral RNA uridylation

Uridylation is not limited to eukaryotic RNAs and has been discovered on many viral RNAs. Using an oligo(dA)-primed RT-PCR combined with Sanger sequencing, Huo *et al.*, characterized the 3' end of viral RNAs from viruses infecting fungi, plants and animals (Huo et al., 2016). They conducted the analysis on the positive-strand RNA viruses TMV (*Virgaviridae*), CMV (*Bromoviridae*), TRV (*Virgaviridae*), ORSV (*Virgaviridae*), TCV (*Tombusviridae*), PVX (*Alphaflexviridae*), PVY (*Potyviridae*), infecting plants and which genome terminates either with a TLS or other structures or a poly(A) tail, as well as the fungus and animal infecting viruses FgHV2 (*Hypoviridae*), PEDV (*Coronaviridae*) and PRRSV (*Arteriviridae*), which genome are polyadenylated. Uridine-rich tails have been detected at the 3' end of all these viral RNAs. Next, they pursued their exploration on negative-strand RNA viruses with the influenza A virus (IAV, *Orthomyxoviridae*) and the rice stripe virus (*Tenuivirus*, unassigned family), infecting animals and plants, respectively. Their segmented genome of negative polarity is transcribed into positive strand RNAs from which the viral proteins are translated. In this study, uridylation has been observed at the 3' end of these positive strand RNAs for the genomic segments they analyzed. They have completed the analysis with double-stranded RNA viruses, the *Rice dwarf virus* (RDV, *Reoviridae*) infecting plants, and a mycovirus, the Alternaria longipes dsRNA virus 1 (AlRV, unclassified). Their data show uridylation at the 3' extremities of both (+) and (-) strands for some of RDV segments and on the (+) strand of AlRV genome. Their work opened a new chapter in the study of the functions related to RNA uridylation, that could be involved in viral RNA metabolism.

However, their strategy relies on an oligo(dA)-primed cDNA synthesis, which induces strong bias towards long uridine tails. Yet, additions of 1-3 uridines are the most commonly detected (Chang et al., 2014; Lim et al., 2014; Morgan et al., 2017; Rissland and Norbury, 2009; Scheer et al., 2021; Sement et al., 2013). In addition, they analyzed 3' tailing by Sanger sequencing on a few clones for each viral RNA. So Huo *et al.* might have missed a major proportion of uridine additions due to the strategy they

used. Recent methods have been developed to analyze tailing with more depth, such as TAIL-seq and 3'RACE-seq (Chang et al., 2014; Scheer et al., 2020) and could be used to study with more precision 3' uridylation of various RNAs.

A few years later, the first study investigating the function of viral RNA uridylation in *C. elegans* and mammalian cells was published (Le Pen et al., 2018). By performing a genetic screen in *C. elegans* infected by the Orsay virus (OrV), the authors identified the TUTase CDE-1 as an antiviral factor, that functions independently from RNA interference. OrV RNAs are uridylated at their 3' extremity by CDE-1, as depletion of the protein completely abolishes terminal uridylation on viral RNAs. They propose uridylation as a new antiviral mechanism that could occur on single-stranded viral RNA, while RNA interference would target double-stranded RNAs generated during genome replication. As CDE-1 is a homolog of mammalian TUTases TUT4/7, they also tested their action against the influenza A virus (IAV) in mammalian cells. They only detected terminal uridylation on IAV mRNAs and not on the genomic RNA nor on the complementary RNAs that serves as template to produce more viral genomes. So, they confirmed the results of Huo *et al.* that IAV mRNAs can be uridylated *in vivo*. Moreover, they observed that when the TUTases are knock-down, IAV mRNA and protein levels were higher.

In light of these results, the study of viral RNA uridylation becomes relevant, in particular for plant viruses that are mainly RNA viruses. One study reports an antiviral role of uridylation, which may be related to RNA degradation. As RNA uridylation functions in many different ways in cellular RNA metabolism, it is not excluded that it could have other functions for viral RNAs as well. Our project on viral RNA uridylation in plants aims to investigate the diversity of the molecular mechanisms related to uridylation.

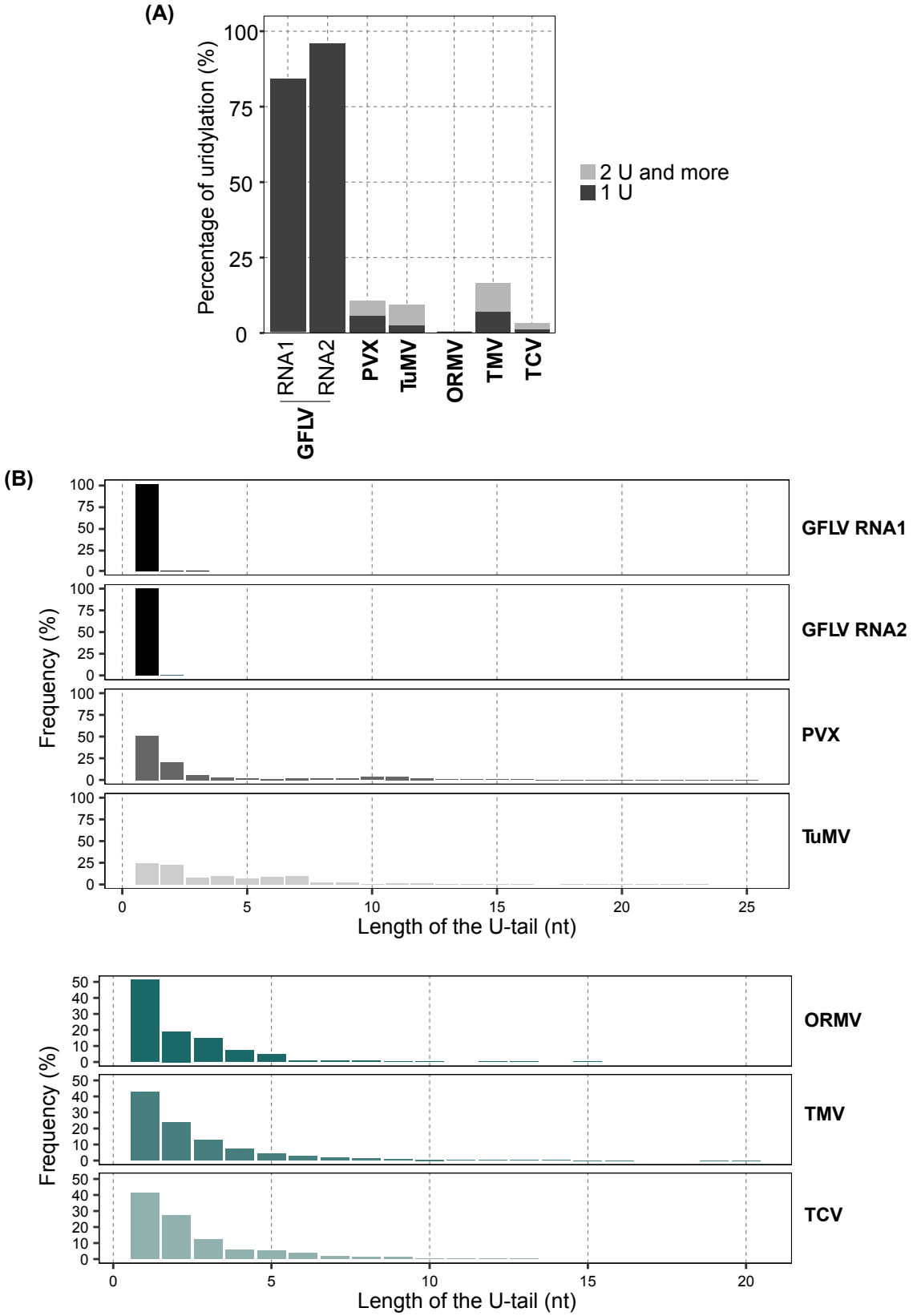


Figure 12: Preliminary data on uridylation patterns of plant viral RNAs. (A) Barplot showing the percentage of uridylation of GFLV, PVX, TuMV, ORMV, TMV and TCV RNAs in one infected plant per virus. The percentage of uridylation is normalized on the total number of reads for each viral RNA. **(B)** Barplot representing the length of the uridine extensions for each viral RNA. The length of the extension is normalized on the number of uridylated reads for each viral RNA. One plant per virus has been analyzed. For polyadenylated RNAs (GFLV, PVX, TuMV), uridines tails on polyadenylated and non-polyadenylated molecules have been considered.

Thesis objectives

In plants, viral infections are an important cause of crop damages worldwide (Sastry and Zitter, 2014). Moreover, the emergence of viral crop diseases and their distribution on the globe could be affected, even enhanced, by climate change (Anderson et al., 2004). Most viruses infecting plants possess an RNA genome. To date, there are 106 phytovirus genera with an RNA genome have been identified, among which 80% are of positive polarity (ICTV, 2022). Given the damages that viruses can cause in crops, it is crucial to better understand the mechanisms related to viral infection and virus-host interaction. Since a few years, uridylation emerges as a new potential regulator of viral infection. One study discovered 3' terminal uridylation on several viral RNAs from viruses infecting fungi, animals and plants (Huo et al., 2016). More recently, Le Pen *et al.* (Le Pen et al., 2018) have highlighted that *C. elegans* TUTase CDE-1 acts as antiviral factor against the Orsay virus. In addition, they have also shown that the mammalian TUTases TUT4 and TUT7 uridylate influenza A virus mRNAs and that it causes their destabilization. The authors of this study propose RNA uridylation as new antiviral mechanism. Yet, mechanistic information related to molecular processes underlying viral RNA uridylation and its consequences remains scarce and to date, there is no study investing uridylation and the related mechanisms for plant viruses.

RNA uridylation targets many types of RNAs, from small RNAs to mRNAs and emerged this last decade as a key player of post-transcriptional regulation in eucaryotes (De Almeida et al., 2018b; Liudkovska and Dziembowski, 2021; Warkocki et al., 2018a; Yu and Kim, 2020; Zigáčková and Vaňáčová, 2018). Our team largely contributed to the investigation of RNA uridylation mechanisms in plants, in particular in the case of mRNAs (Scheer et al., 2021; Sement et al., 2013; Zuber et al., 2016).

In recent years, our team and others have developed the 3'RACE-seq strategy that allows in-depth determination of 3' templated and non-templated nucleotides for target RNAs. This protocol has been optimized for various RNA types, including mRNAs (Scheer et al., 2020) and 5' RISC-cleaved fragments as part of my Master 1 internship project (Zuber et al., 2018).

Before the beginning of my PhD, the team took advantage of this method to analyze 3' end uridylation on viral RNAs from six plant viruses: the grapevine fanleaf virus (GFLV, *Secoviridae*), the potato virus X (PVX, *Alphaflexviridae*), the turnip mosaic virus (TuMV, *Potyviridae*), the oilseed rape mosaic virus (ORMV, *Tobamovirus*), the tobacco mosaic virus (TMV, *Tobamovirus*) and the turnip crinkle virus (TCV, *Betacarmovirus*). All of these viral RNAs are uridylated and the percentage varies from 0.5% for ORMV RNA up to more than 80% for both of GFLV RNAs (Fig. 12A). There are two distinct uridylation patterns: GFLV RNAs are mono-uridylated whereas the others have uridines extensions between 1 and 30 Us, with a majority of 1-3 U-tails (Fig. 12B). Interestingly, to our knowledge, GFLV RNA uridylation profile is unique and has never been observed for plant mRNAs. These preliminary results indicate that uridylation profiles are diverse on plant viral RNAs. In eukaryotes, uridylation involves many different factors to regulate complex molecular process related to both coding and non-coding RNAs. For viral RNAs, the various uridylation patterns also tend to indicate that different molecular actors and mechanisms are in place.

In this context, the aim of my thesis project was to contribute to a better understanding of molecular processes linked to plant viral RNA uridylation. In particular, key questions arise from the actual state of the art on plant viral RNA uridylation:

- What is the diversity of uridylation patterns present on plant viral RNAs?
- What are the enzymes responsible for the uridine additions, are they host or viral factors?
- What is the impact of uridylation on the viral RNA stability?
- What is the role of RNA uridylation in the host-virus interactions?

The first objective of my PhD work was to investigate the diversity of uridylation patterns in plant viral RNAs. Our team initiated a collaboration with the group of Olivier Lemaire at the INRAE of Colmar (France) that is specialized in the study of grapevine-infecting viruses. I was actively involved in this collaboration for the organization and the sample collection, as well as in performing the experiments and analyzing results. Altogether, we analyzed by 3'RACE-seq uridylation patterns on more than thirty plant viral RNAs, that are representatives of the main ss(+) RNA phytovirus families, including a focus on *Secoviridae* species to which GFLV belongs.

The second objective of my thesis was to identify factors involved in viral RNA uridylation. For this objective, I focused on three model viruses – GFLV, TuMV and TCV – that infect and replicate in our model plant *Arabidopsis thaliana*. I evaluated the impact of the inactivation of URT1 and HESO1, the TUTases identified in Arabidopsis to date, on the uridylation pattern of the viral RNAs from these viruses. My results reveal that host and possibly viral factors are involved and that at least two mechanisms are in place.

The last objective of my work was to initiate the study of the impact of uridylation on the viral metabolism. I first analyzed the impact of an absence or of an overexpression of Arabidopsis TUTases on viral accumulation *in planta*. I generated transgenic plants overexpressing HESO1 while URT1 overexpressing lines were already available in the team. In parallel, I started the analysis of viral RNA accumulation and uridylation in plants mutated for cytoplasmic RNA decay factors.

The results that I collected during my PhD work constitute a resource on plant viral RNA uridylation, that is necessary to better apprehend the potential pro- and antiviral roles of uridylation in the viral metabolism. My data also provide insights on the actors involved in viral RNA uridylation in plants and the project paves the way to better understand the potential link between uridylation and degradation for plant viral RNAs.

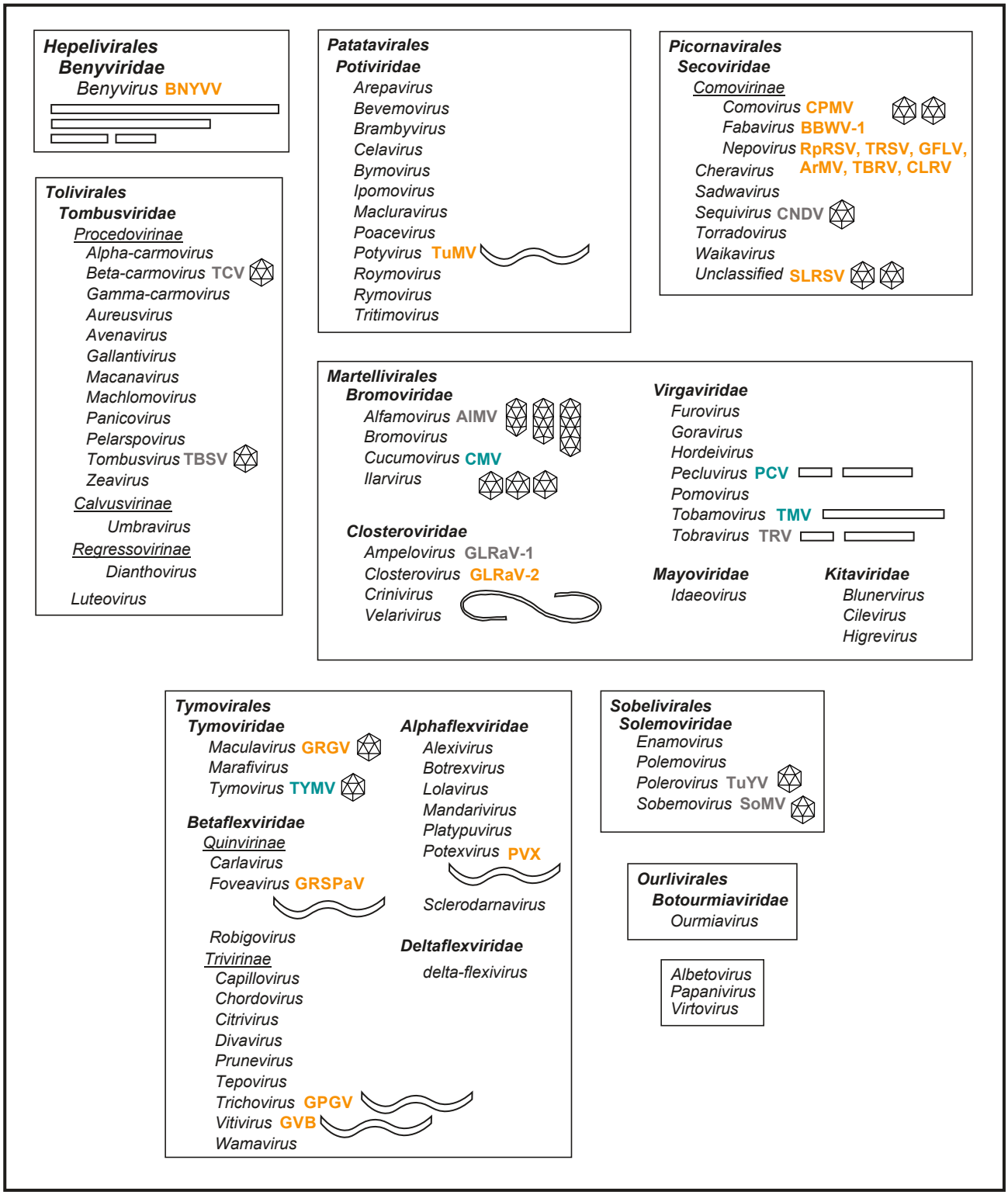


Figure 13: Overview of the ss(+) RNA phytoviral families according to recent taxonomy (ICTV 2021). Viral species selected for 3'RACE-seq analysis are written next to their genus. The color code indicates the 3' RNA feature: orange (poly(A) or A-rich tail), green (known tRNA-like structure) and grey (not polyadenylated, with RNA structures or unknown). The number and morphology of virions are schematized (icosahedral, flexous and rod-shaped particles).

Presentation of the viruses used in this study

The first objective of the project was to investigate the diversity of uridylation patterns among viral RNAs. In this chapter, I will describe the viruses that we used for this study. For this part of the project, we cooperated closely with the team of Olivier Lemaire (team Virology-Vectors, INRAE Colmar, France) who is specialized in the study of grapevine-infecting viruses. To obtain infectious material for some of our viruses of interest, we also had help of other colleagues: Salah Bouzoubaah, David Gilmer, Anthony Gobert, Manfred Heinlein, Christophe Ritzenthaler, and Véronique Ziegler-Graff (IBMP, Fance); Corinne Schmitt-Keichinger, Véronique Brault, and Claire Villeroy (SVQV, INRAE Colmar, France); Marc Fuchs (Cornell University, USA); Jean-Sébastien Reynard (Virology-Phytoplasmalogy Laboratory, Agroscope, Switzerland).

We chose viruses representing the main families of ss(+) RNA phytoviruses. Twenty-nine viruses covering 11 out of 15 families of plant viruses have been chosen: 1) *Benyviridae* (one species), 2) *Bromoviridae* (two species), 3) *Closteroviridae* (two species), 4) *Virgaviridae* (three species), 5) *Potyviridae* (one species), 6) *Secoviridae* (ten species), 7) *Solemoviridae* (two species), 8) *Tombusviridae* (two species), 9) *Alphaflexviridae* (one species), 10) *Betaflexviridae* (three species) and 11) *Tymoviridae* (two species). In line with the research interests of our collaborators for this project, O. Lemaire's group, more than half of the viruses that we analyzed infect grapevine. The Fig. 13 presents an overview of the phytoviral families related to the taxonomy and the viruses that we analyzed are indicated.

We used viruses with different 5' and 3' RNA features that are recapitulated in Table 3. Thirteen viral genomes have a 5' cap structure and thirteen have a 5' VPg, while three does not have either of these structures. On the 3' extremity, seventeen viral genomes are polyadenylated, four possess a tRNA-like structure and eight have other types of structures.

Order	Family	Subfamily	Genus	5' feature	3' feature	Viral species	Reference
Hepelivirales	*Benyviridae*	-	*Benyvirus*	cap	polyA	BNYVV	Gilmer & Ratti, 2017
Martellivirales	*Bromoviridae*	-	*Alfamovirus*	cap	no polyA	AlMV	Olsthoorn 1999
Martellivirales	*Bromoviridae*	-	*Cucumovirus*	cap	TLS	CMV	Ahlquist 1981
Martellivirales	*Closteroviridae*	-	*Ampelovirus*	cap	no polyA	GLRaV-1	Fuchs 2020
Martellivirales	*Closteroviridae*	-	*Closterovirus*	cap	A-rich tails	GLRaV-2	Zhu 1998
Martellivirales	*Virgaviridae*	-	*Pecluvirus*	cap	TLS	PCV	Herzog 1994
Martellivirales	*Virgaviridae*	-	*Tobamovirus*	cap	TLS	TMV	Goelet 1982
Martellivirales	*Virgaviridae*	-	*Tobravirus*	cap	Structure resembling TLS (TRV)	TRV	van Belkum 1987
Patatavirales	*Potyviridae*	-	*Potyvirus*	VPg	polyA	TuMV	Nicolas and Laliberté 1992
Picornavirales	*Secoviridae*	-	*Sequivirus*	VPg	no polyA	CNDV	This work for CNDV
Picornavirales	*Secoviridae*	*Comovirinae*	*Comovirus*	VPg	polyA	CPMV	Thompson 2017
Picornavirales	*Secoviridae*	*Comovirinae*	*Fabavirus*	VPg	polyA	BBWV-1	Thompson 2017
Picornavirales	*Secoviridae*	*Comovirinae*	*Nepovirus* A	VPg	polyA	RpRSV	Thompson 2017
Picornavirales	*Secoviridae*	*Comovirinae*	*Nepovirus* A	VPg	polyA	TRSV	Thompson 2017
Picornavirales	*Secoviridae*	*Comovirinae*	*Nepovirus* A	VPg	polyA	GFLV	Ritznethaler 1991, Serghini 1990
Picornavirales	*Secoviridae*	*Comovirinae*	*Nepovirus* A	VPg	polyA	ArMV	Thompson 2017
Picornavirales	*Secoviridae*	*Comovirinae*	*Nepovirus* B	VPg	polyA	TBRV	Thompson 2017
Picornavirales	*Secoviridae*	*Comovirinae*	*Nepovirus* C	VPg	polyA	CLRV	Thompson 2017
Picornavirales	*Secoviridae*	-	Unassigned	VPg	polyA	SLRSV	Thompson 2017
Sobelivirales	*Solemoviridae*	-	*Polerovirus*	VPg	no polyA	TuYV	Sõmera 2021
Sobelivirales	*Solemoviridae*	-	*Sobemovirus*	VPg	no polyA	SoMV	Sõmera 2021
Tolivirales	*Tombusviridae*	*Procedovirinae*	*Betacarmovirus*	-	no polyA	TCV	Simon 2015
Tolivirales	*Tombusviridae*	*Procedovirinae*	*Tombusvirus*	-	no polyA	TBSV	Fabian 2003
Tymovirales	*Alphaflexviridae*	-	*Potexvirus*	cap	polyA	PVX	Huisman 1988
Tymovirales	*Betaflexviridae*	*Quinvirinae*	*Foveavirus*	cap	polyA	GRSPaV	Adams 2009

Tymovirales	*Betaflexviridae*	*Trivirinae*	*Trichovirus*	-	polyA	GPGV	Adams 2009
Tymovirales	*Betaflexviridae*	*Trivirinae*	*Vitivirus*	cap	polyA	GVB	Adams 2009
Tymovirales	*Tymoviridae*	-	*Maculavirus*	cap	polyA	GRGV	Dreher 2009
Tymovirales	*Tymoviridae*	-	*Tymovirus*	cap	TLS	TYMV	Dreher 2010

Table 3: 5' and 3' terminal RNA features of the viruses used in this study. The classification from order to genus is indicated for each species. The viral species are written in their acronym form: BNYVV = Beet necrotic yellow vein virus, AlMV = *Alfalfa mosaic virus*, CMV = *Cucumber mosaic virus*, GLRaV-1 = *Grapevine leafroll virus 1*, GLRaV-2 = *Grapevine leafroll virus 2*, PCV = *Peanut clump virus*, TMV = *Tobacco mosaic virus*, TRV = *Tobacco rattle virus*, TuMV = *Turnip mosaic virus*, CNDV = *Carrot necrotic dieback virus*, CPMV = *Cowpea mosaic virus*, BBWV-1 = *Broad bean wilt virus 1*, RpRSV = *Raspberry ringspot virus*, TRSV = *Tobacco ringspot virus*, GFLV = *Grapevine fanleaf virus*, ArMV = *Arabis mosaic virus*, TBRV = *Tomato black ring virus*, CLRV = *Cherry leaf roll virus*, SLRSV = *Strawberry latent ringspot virus*, TuYV = *Turnip yellows virus*, SoMV = *Sowbane mosaic virus*, TCV = *Turnip crinkle virus*, TBSV = *Tomato bushy stunt virus*, PVX = *Potato virus X*, GRSPaV = *Grapevine rupestris stem pitting-associated virus*, GPGV = *Grapevine pinot gris virus*, GVB = *Grapevine virus B*, GRGV = *Grapevine red globe virus*, TYMV = *Turnip yellow mosaic virus*. The color code for the type of 3' extremity is similar to Fig. 13.

Preliminary data obtained in our team has revealed a singular uridylation pattern for GFLV, member of the *Secoviridae* (Fig. 12A), so we decided to extent the analysis to more species from this family. In addition to the analysis of the uridylation landscape among phytoviral RNAs, we wanted to enlighten the actors of viral RNA uridylation, in particular whether host TUTases could be involved. So, to carry out the analysis in wild-type and plants mutated for these factors, we chose three viruses that can replicate in our model plant *Arabidopsis thaliana*: the *Turnip crinkle virus* (*Tombusviridae*), the *Turnip mosaic virus* (*Potyviridae*) and the *Grapevine fanleaf virus* (*Secoviridae*), on which I will give more details.

a. Viruses with a polyadenylated genome

Benyviridae

We studied one viral species from the *Benyviridae*, the beet necrotic yellow vein virus (BNYVV). *Benyvirus* is the only genus within this family and it regroups multipartite plant viruses with non-enveloped rod-shaped viral particles. The genome is segmented and each segment possesses a 5' m^7G cap structure and a 3' poly(A) tail. BNYVV is a typical member of this family and its genome comprises four to five segments, depending on the isolate. RNA1 codes for the replicase proteins while RNA2 encoded-ORFs produce the capsid protein, the triple gene block proteins (TGBp) involved in cell-to-cell movement, and a suppressor of silencing (Andika et al., 2012; Chiba et al., 2013; Gilmer and Ratti, 2017). Three subgenomic RNAs are produced from RNA2 to translate the TGBp and the suppressor of silencing. Of note, RNA1 and RNA2 are sufficient for replication under experimental conditions. RNA3 presence is associated with symptoms and leads to the accumulation of noncoding RNA3 (Peltier et al., 2012). RNA4-encoded protein is associated with symptoms expression and virus transmission (D'Alonzo et al., 2012). Some isolates also have a fifth RNA, RNA5, which also leads to the accumulation of a noncoding RNA (Peltier et al., 2012). Benyviruses are transmitted by root-infecting single-cellular eukaryotes from the *Cercozoa* phylum, causing the "rhizomania" disease in their natural host *Beta vulgaris*. Under experimental conditions, other species can be infected such as *Chenopodium quinoa*, locally, and *Beta macrocarpa, Nicotiana benthamiana* and *Spinacea olacera*, systemically (Gilmer and Ratti, 2017).

Potyviridae

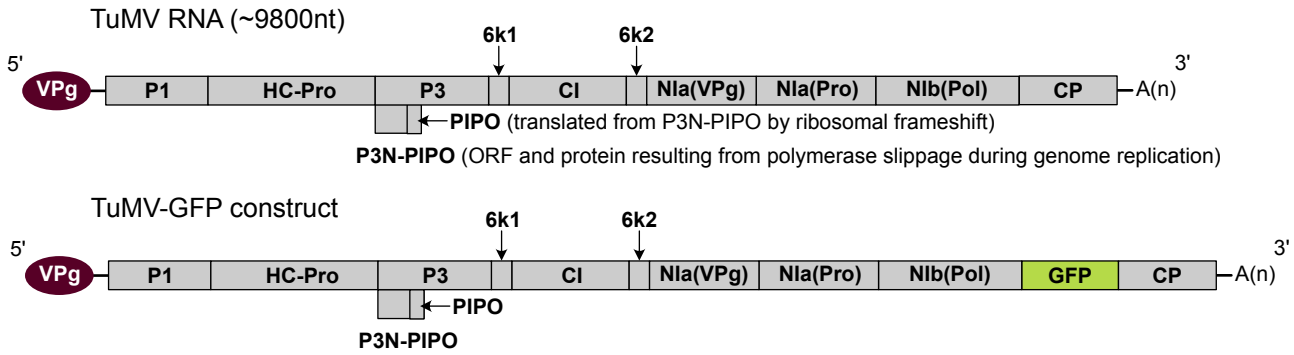
General informations on TuMV

The turnip mosaic virus (TuMV) is model viral species broadly studied. It is classified in the *Potyviridae* family which members have a mono- or bipartite genome contained in non-enveloped flexuous filamentous particles (Inoue-Nagata et al., 2022).

This virus was first isolated from *Brassica rapa* in the early 1920s (Schultz, 1921). TuMV has a huge host range which includes more than 300 species, distributed among 43 plant families (Nellist et al., 2022). It can also systematically infect the model plant Arabidopsis, which is very useful to study factors involved in the viral cycle and to find resistance genes. In nature, potyviruses are transmitted by

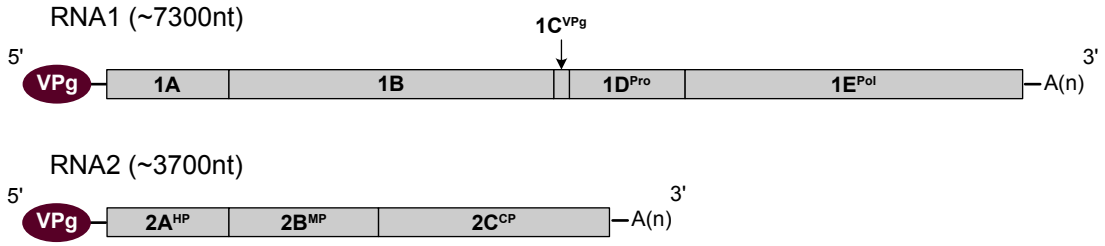
(A) Overview of the genome organisation of the ***Turnip mosaic virus*** (TuMV) and TuMV-GFP

P1 = protein 1, HC-Pro = helper-component protease, P3 = protein 3, 6k1 and 6k2 = 6 kDa peptide 1 and 2, respectively, CI = cylindrical inclusion protein, VPg = viral protein genome-linked, NIa = nuclear inclusion A-protease, NIb = nuclear inclusion B (RNA dependent RNA polymerase), CP = capsid protein, P3N-PIPO = protein 3 N-terminal pretty interesting *Potyvirus* ORF (PIPO), GFP = green fluorescent protein.
All viral proteins are generated by proteolytic cleavage from a polyprotein, except for P3N-PIPO and PIPO.



(B) Overview of the genome organisation of the ***Grapevine fanleaf virus*** (GFLV)

1A = protein 1A, 1B = protein 1B (helicase), 1C = viral protein genome-linked, 1D = protease, 1E = RNA dependent RNA polymerase, 2A = homing protein, 2B = movement protein, 2C = capsid protein.
All viral proteins are generated by proteolytic cleavage from polyproteins resulting from RNA1 and RNA2, respectively.



(C) Overview of the genome organisation of the ***Turnip crinkle virus*** (TCV)

P28 leaky termination site (indicated by a dotted line) allows ribosomal readtthrough ("RT") to produce the RNA dependent RNA polymerase. P8 and P9 are translated from a subgenomic RNA into the movement protein. P38 codes for the capsid protein and is translated from a second subgenomic RNA.

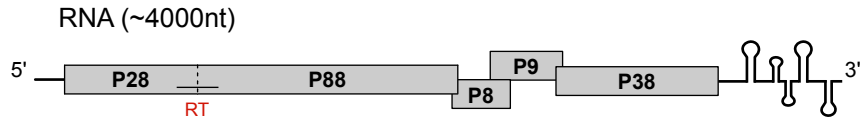


Figure 14: Genome organization of the three model viruses of this study: (A) *Grapevine fanleaf virus,* **(B)** *Turnip mosaic virus* and the clone expressing the GFP, TuMV-GFP, and **(C)** *Turnip crinkle virus.* The names of the viral proteins are indicated in the corresponding ORFs. TuMV and GFLV genomes are linked to a viral protein at their 5' extremity, the genome-linked viral protein (VPg) and they have a 3' poly(A) tail, symbolized by A(n). TCV RNA 3'UTR is contains RNA structures.

aphids in a non-persistent and non-circulative manner. TuMV can be transmitted by many species of aphids, in particular *Myzus persicae* (Edwardson and Christie, 1986; Nellist et al., 2022).

The viral replication takes place in viral factories within ER membrane-derived vesicles (Jin et al., 2018). The model for TuMV replication and translation determined by confocal microscopy proposes that once the viral RNA genome is released into the cytoplasm, the translation of the viral proteins starts on ER-associated ribosomes. Then, newly synthetized viral proteins initiate the invagination of the ER-membrane, thereby entrapping viral proteins and host translation machinery in vesicles. Translation coupled to replication continues within the vesicles (Grangeon et al., 2010).

Genomic organization of TuMV

TuMV complete genome is about 9.8 kb long. The 5'UTR is predicted to fold into a hairpin and it has a 200 nt long 3'UTR followed by a poly(A) tail (Nicolas and Laliberte, 1992). A VPg is covalently linked to the 5' extremity. The genome contains one single ORF that encodes for a polyprotein (Fig. 14A) from which the viral proteins are generated by self-cleavage as well as a small protein produced by polymerase slipping (Revers and García, 2015).

The 5' proximal protein 1 is a protease that results from self-cleavage by a serine protease domain, followed by the helper component-protease HC-Pro, self-cleaved by a cysteine protease domain. HC-Pro is important in suppression of gene silencing and vector transmission. Protein 3 plays a role in virus replication and movement. The two small peptides 6K1 and 6K2 are required for efficient replication, with 6K2 being a transmembrane protein that anchors the replication complex to the ER membrane. The 6K2-induced vesicles can move to the bordering cells (Grangeon et al., 2013). The cylindrical inclusion protein (CI) has been named after its ability to accumulate in inclusion bodies within the cytoplasm and plasmodesmata of infected cells and it has a helicase activity. The VPg originates from the first region of the nuclear inclusion a (NIa) protein, while the second region of this protein generates a serine-like cysteine protease that cleaves at most of the sites within the potyvirus polyprotein and accumulates in inclusions in the nucleus. Another protein accumulates in inclusion within the nucleus, the viral RdRp, or NIb. The 3' proximal protein codes for the capsid protein that is also involved in virus movement (Revers and García, 2015). In addition, the eleventh viral protein is produced from the short pretty interesting *Potyvirus* ORF (PIPO) located within the protein 3 reading frame. PIPO is expressed by polymerase slippage and is required for cell-to-cell movement (Chung et al., 2008). PIPO expressed with CI are enough to induce viral movement to the adjacent cells through the plasmodesmata (Movahed et al., 2017).

TuMV RNA and interactions with host factors

TuMV replication requires host PABP (Beauchemin and Laliberté, 2007; Dufresne et al., 2008a). In addition to PABP, the heat shock cognate 70 3 (Hsc70-3) protein interacts with TuMV RdRp *in vitro* and is relocalized from the nucleus into viral factories (Dufresne et al., 2008b). Decapping factors are also involved in TuMV replication (Zuo et al., 2022). More details about TuMV RNA and its interactions with host factors, in particular factors related to RNA degradation are provided in Chapter 4.2 of the Introduction. To my knowledge, there is no information on the synthesis of TuMV RNA poly(A) tail.

Secoviridae

Overview of the Secoviridae

Viral species from the *Secoviridae* family have non-enveloped icosahedral viral particles containing a mono- or bipartite RNA genome. *Secoviridae* species mostly infect dicotyledonous plants and are transmitted by nematodes or insects. This viral family is related to *Picornaviridae*, important pathogens infecting mammals. The viral genome has a 5' VPg and is polyadenylated (Thompson et al., 2017). However, data on two members from the *Sequivirus* genus suggest that sequivirus genomes are not polyadenylated. The RNA of the parsnip yellows fleck virus is the only genome from this family that is completely sequenced. It possesses a longer 3'UTR than other *Secoviridae* which the last 121 nt that are predicted to fold into a stem loop (Turnbull-Ross et al., 1992). Studies on another sequivirus, the lettuce mottle virus have demonstrated that it cannot be amplified from oligo(dT)-primed cDNA, which could indicate that the viral genome is not polyadenylated (Jadão et al., 2007; Menzel and Vetten, 2008).

In this study, we used the carrot necrotic dieback virus (CNDV), assigned to the *Sequivirus* genus. We also used the cowpea mosaic virus (CPMV, *Comovirus*) the broad bean wilt virus 1 (BBWV-1, *Fabavirus*), the strawberry latent ringspot virus (SLRSV, *Stralarivirus*) and we analyzed the following nepoviruses: the raspberry ringspot virus (RpRSV), the tobacco ringspot virus (TRSV), the grapevine fanleaf virus (GFLV), the arabis mosaic virus (ArMV), the tomato black ring virus (TBRV) and the cherry leaf roll virus (CLRV). Nepoviruses are divided into three subgroups: A, B and C, according to the size of the RNA2 and the viral particle composition. RpRSV, TRSV, GFLV and ArMV are from subgroup A, TBRV from subgroup B and CLRV from subgroup C.

The presence of a small VPg (2-4 kDa) at the 5' extremity of viral RNA has been demonstrated for some comoviruses and nepoviruses as well as for SLRSV. The viral RNAs code for polyproteins from which viral proteins are cleaved by viral 3C-like proteinases. An exception to this is the RNA2 of comoviruses which contains two ORFs, resulting in the production of two polyproteins. RNA1-encoded polyprotein codes for a putative helicase, a 3C-like proteinase and the viral polymerase. Structural proteins and factors required for cell-to-cell movement are generated from RNA2-encoded polyprotein. For CNDV, which has a monopartite genome, the structural proteins are located upstream of the replicase proteins. Each genus and the different subgroups of nepoviruses are defined by the cleavage site specificity of the proteinase. The genera are also distinguishable according to their transmission vector. Comoviruses are transmitted by beetles, fabaviruses by aphids, nepoviruses by nematodes and sequiviruses by aphids, with the help of a helper virus. To date, the mode of transmission of SLRSV and its genome organization are unknown (Thompson et al., 2017).

General informations on GFLV

The grapevine fanleaf virus (GFLV) and other nepoviruses are the agents of a grapevine degenerative disease, the grapevine fanleaf disease. The first reports of this disease in Europe have been published 150 years ago. The main symptoms are cane malformations and yellow mosaic on the leaves. On the contrary to other viruses from this family, GFLV has a narrow host range in nature restricted to *Vitis* species (Meng et al., 2017) into which it is transmitted via the nematode *Xiphinema index* when it is feeding on the roots (Andret-Link et al., 2004). Under experimental conditions, it can be

mechanically inoculated onto *C. amaranticolor*, *C. quinoa*, *N. benthamiana* and *A. thaliana*. Of note, it can systematically infect Arabidopsis but the infection remains symptomless.

Genomic organization of GFLV and replication

GFLV genome consists of two polyadenylated RNA molecules (RNA1 and RNA2). RNA1 possesses one ORF translated into a polyprotein that generates five viral proteins by cleavage (Fig. 14B): a proteinase cofactor, a helicase, the VPg, the proteinase and the viral RdRp (Margis et al., 1994; Pinck et al., 1991; Ritzenthaler et al., 1991). RNA2 encodes a polyprotein that is cleaved into three proteins (Fig. 14B): the homing protein that is involved in viral RNA replication, the movement protein and the capsid protein (Gaire et al., 1999; Margis et al., 1993; Ritzenthaler et al., 1995; Serghini et al., 1990). The 3'UTRs of both RNAs present a high sequence homology (Ritzenthaler et al., 1991). GFLV also contains satellite RNAs. They have been reported to be associated with GFLV infection, in particular for the F13 isolate (Pinck et al., 1988).

GFLV replication takes place in ER membrane-derived vesicles (Ritzenthaler et al., 2002). GFLV belongs to the *Secoviridae* family, related to *Picornaviridae*, and its genome has a small VPg that is covalently linked to the 5' of the genome by a phosphodiester bond between a serine and the first nucleotide (Pinck et al., 1991). It is not known whether the replication process could be protein-primed similarly to what was shown for the poliovirus VPg uridylylation. To my knowledge, there is no experimental work investigating GFLV genome replication nor on the synthesis of the poly(A) tail.

Alphaflexviridae

Viral species with relatively small flexuous filamentous virions are regrouped within this family which is divided into six genera. Their genome has a 5' cap structure and is polyadenylated. The members of this family infect plants or fungi (Kreuze et al., 2020). The potato virus X (PVX) is a *Potexvirus* belonging to this family and one of the most extensively studied phytovirus (Scholthof et al., 2011).

Potexviruses infects herbaceous plants and no vectors have been reported so far. PVX genome has five ORFs. The most 5' proximal one codes for the RdRp. It is followed by the triple gene block (TGB) of three overlapping ORFs that encode movement proteins and that result from the translation of sgRNAs. TGBp1 is also a suppressor of silencing. The capsid protein is generated by the translation of the last ORF from a sgRNA (Kreuze et al., 2020).

Betaflexviridae

This family of plant-infecting viruses regroups more than 100 species classified within fifteen genera. The virions are flexuous filaments, longer than those of *Alphaflexviridae*. Accordingly, their genome is larger (Adams et al., 2009). In our work, three genera of *Betaflexviridae* are represented: *Foveavirus* with the grapevine rupestris stem pitting-associated virus (GRSPaV), *Trichovirus* with the grapevine pinot gris virus (GPGV) and *Vitivirus* with the grapevine virus B. Both GRSPaV and GVB monopartite genomes are presumably capped and possess a poly(A) tail whereas GPGV monopartite genome is polyadenylated but does not have a 5' cap structure (Adams et al., 2009).

Foveavirus monopartite genome possesses 5 ORFs and the capsid protein is larger than most members of the *Betaflexviridae*, for example GRSPaV capsid protein measures 28 kDa. GRSPaV host range is restricted to grapevine species, in which it causes the rugose wood disease, and no vector is described yet. The genome codes for six ORFs. The first ORF encodes a putative polyprotein involved in replication that contains several domains including methyltransferase, helicase and RdRp domains. Downstream, three ORFs codes for the TGB, likely to be involved in cell-to-cell movement based on the work on PVX. The capsid protein is encoded by the ORF located after the TGB. A sixth ORF overlapping the CP-encoding ORF has been discovered but its mode of translation and potential function remain unknown (Meng et al., 2017).

GPGV belongs to the *Trichovirus* genus and the typical monopartite genome encodes three or four ORFs, including a movement protein that is quite large in size (Adams et al., 2009). GPGV infects grapevine and is associated with grapevine leaf mottling and deformation disease, but it can also be found in symptomless grapevines. Of note, two herbaceous hosts have also been reported. GPGV is transmitted by a mite species (Meng et al., 2017). The genome has three ORFs, coding for a putative replicase, movement and capsid proteins (Giampetruzzi et al., 2012). So far, there is no experimental data on the strategy of translation and replication of GPGV, but it is likely to depend on polyprotein processing for the first ORF and on the production of sgRNAs for the others (Meng et al., 2017).

Grapevine vitiviruses regroup several viral species associated with the rugose wood disease. GVB is a *Vitivirus* that has been discovered in the 1990s. In nature, this virus only infects grapevine via mealybug transmission and under experimental conditions, it can infect herbaceous hosts. GVB monopartite genome contains five ORFs and has a polyadenylation signal in the 3'UTR. The first ORF contains, like in foveaviruses and trichoviruses, methyltransferase, helicase and RdRp domains. No protease domain was found in vitivirus genomes and there is no experimental data on the strategy of translation. The second ORF is genus-specific and codes for a polypeptide that could be required for vector transmission. ORF3 and ORF4 code for the large movement protein and the capsid protein, respectively. The 3' proximal ORF translation generates a small RNA-binding protein of unknown function (Meng et al., 2017).

Maculavirus (Tymoviridae)

The *Tymoviridae* viral family regroups species with non-enveloped isometric particles. The genomic RNA has a 5' cap structure. Most species possess a TLS (genus *Tymovirus*) at the 3' extremity of their genome but some have instead a poly(A) tail (genera *Maculavirus* and *Marafivirus*) (Dreher et al., 2009). In this work, we used a grapevine-infecting virus, the grapevine red globe virus (GRGV) which belongs to the *Maculavirus* genus. GRGV mechanism of transmission remains unknown. Its genome is partially sequenced and three ORFs were discovered: the C-terminal part of the first large ORF contains the RdRp domain, the second ORF encodes the capsid protein and the third for a 17-kDa protein of unknown function (Meng et al., 2017).

b. Viruses with a non-polyadenylated genome

Bromoviridae

Bromoviridae species infect plants and have a tri-segmented genome, encapsidated in separate non-enveloped virions (Bujarski et al., 2019). Among the six genera regrouped in the family, we studied one alfamovirus, the alfalfa mosaic virus (AlMV) and one cucumovirus, the cucumber mosaic virus (CMV). CMV belongs to the 'top 10' plant viruses of scientific and economic importance (Scholthof et al., 2011). Alfamovirus virions are bacilliform while the virions of cucumoviruses is spherical.

In nature, they both have a broad host range including alfalfa, bananas, cucurbits and tomatoes (Bujarski et al., 2019). Under experimental conditions, these viruses can also replicate in *A. thaliana* and *N. benthamiana* (Balasubramaniam et al., 2006). There are more than 70 species of aphids that are reported to transmit CMV (Roossinck, 2001). AlMV is also transmitted by aphids, at least fourteen species have been identified, in particular *Myzus periscae* (Swenson, 1952).

The three genomic segments have a 5' cap structure but they are not polyadenylated. Cucumoviruses have a tRNA-like structure at the 3' extremity that can be aminoacylated with tyrosine and the 3'UTR of alfamoviruses folds into a structure resembling a TLS. The TLS structure based on the last 190 bases of CMV RNAs has been proposed, along with other members of the *Bromoviridae* (Ahlquist et al., 1981; Lück et al., 1996). In the case of AlMV, the 3'UTR folds into two conformations that are involved in the switch between translation and replication of the genome. Although one of this conformation involves a pseudoknot like other members of the *Bromoviridae*, it is a poor substrate for the CCA-adding enzyme *in vitro* (Olsthoorn et al., 1999).

RNA1 and RNA2 of alfamoviruses code for replicase proteins and RNA3 for a movement protein and the capsid protein, the latter being translated from a subgenomic RNA that is also capped. Cucumoviruses genome organization is the same and the RNA2 also contains ORF 2b which codes for a silencing suppressor and cell-to-cell movement protein translated from a capped subgenomic RNA (Bujarski et al., 2019).

Closteroviridae

This family regroups plant viruses that have a long helicoidal filamentous non-enveloped particles and a large genome (13,000 to 19,000 nt), that can be segmented. Our candidate viruses the grapevine leafroll-associated virus 1 (GLRaV-1) and the grapevine leafroll-associated virus 2 (GLRaV-2) belong to two of the four genera, the *Ampelovirus* and the *Closterovirus* genera, respectively.

Their monopartite genome likely harbors a 5' cap structure and the viral RNAs were described as not polyadenylated (Fuchs et al., 2020). However, the genomic sequences published for GLRaV-2 reveals a poly(A) tract, as previously described by Zhu and colleagues (Zhu et al., 1998).

At least eleven GLRaVs are responsible for the grapevine leafroll-associated disease in their natural host grapevine and they all belong to the *Closteroviridae* family. GLRaV-1 and GLRaV-2 have been discovered in naturally infected vineyards in France (Meng et al., 2017).

In ampeloviruses, the genome has twelve ORFs, including two coding for the replicase proteins, two coding for the capsid protein and the minor capsid protein, respectively, and proteins involved in

cell-to-cell movement as well as a suppressor of silencing (Fuchs et al., 2020; Zhang et al., 2020a). Cell-to-cell movement factors, structural proteins and suppressors of silencing are translated from eleven distinct sgRNAs. GLRaV-1 has one the largest genomes among *Closteroviridae* (18,000 nt), which encodes two copies of the minor capsid protein. It is transmitted by several genera of mealybugs and soft scale insects.

The genome of closteroviruses is smaller (14,500-19,300 nt) and encodes a similar set of proteins, some of which also being translated from sgRNAs. The range of vectors is variable between closteroviruses and is unknown for GLRaV-2 yet (Fuchs et al., 2020).

Virgaviridae

This family regroups plant viruses with a rod-shaped virion containing a genomic RNA that is multipartite except for tobamoviruses. The viral particles are non-enveloped helical rods. The genome possesses a 5' cap structure and an aminoacylable TLS at its 3' extremity, except for members of the *Tobravirus* genus, as it has been shown for the tobacco rattle virus (Adams et al., 2017; van Belkum et al., 1987).

The peanut clump virus (PCV) belongs to the *Pecluvirus* genus that contains soil-borne viruses with a bipartite genome. Pecluviruses are transmitted by a unicellular eukaryote from the *Plasmodiophoridae* family (*Cercozoa*). PCV natural hosts include *Arachis hypogea* (peanut) and *Sorghum arundinaceum* (common wild sorghum).

Its TLS can be esterified with valine and the structure has been resolved. Compared to other viral TLSs, PCV TLS has an insertion of about 40 nt which does not prevent it to fold into a L-shaped structure (Dreher, 2010; Goodwin and Dreher, 1998). RNA1 codes for the replicase proteins and for a suppressor of silencing that is translated from a subgenomic RNA. RNA2 encodes for the capsid protein, a protein involved in the viral transmission and triple gene block proteins required for the cell-to-cell movement, probably expressed from subgenomic RNAs. Experimental hosts include *C. quinoa, C. amaranticolor* (both local lesions hosts) and *N. benthamiana* (systemic) (Adams et al., 2017).

The tobacco mosaic virus (TMV) is one of the 'top 10' plant viruses of scientific and economic importance (Scholthof et al., 2011) and a typical member of the *Tobamovirus* genus. On the contrary to many phytoviruses, it is not vector-transmissible.

Its genome is monopartite and can be aminoacylated with histidine (Adams et al., 2017). The structure of the TLS has been resolved by experimental data and predictions (Dreher, 2010; Rietveld et al., 1984). The last 105 nt are structured into three hairpins that fold into a L-shaped structure. TMV genome encodes for replicase proteins, that were proposed to function as suppressor of silencing (Ding et al., 2004), movement and capsid proteins, expressed from distinct sgRNAs both capped an ending with a TLS (Adams et al., 2017).

The tobacco rattle virus (TRV) is affiliated to tobraviruses and has a bipartite genome. It is transmitted by nematodes. In nature, TRV has a broad host range including monocotyledonous and dicotyledonous plants like narcissus and potato, respectively (Adams et al., 2017).

Although its 3' extremity can be acylated *in vitro* like a TLS, but at a lesser extent, it cannot be aminoacylated so it is not *de facto* considered as a TLS (van Belkum et al., 1987). Experimental data

show that the structure shares features with tobamoviral and tymoviral RNAs, but some domains are longer or shorter and a domain mimicking the anticodon arm has not been found. RNA1 codes for the replicase proteins, a protein involved in intercellular movement and suppressor of silencing, both translated from sgRNAs. RNA2 codes for the capsid protein translated from a sgRNA and for factors involved in nematode transmission, which translation mechanism remains unknown (Adams et al., 2017).

Solemoviridae

These plant viruses have an icosahedral particle containing a monopartite genome that is not polyadenylated (Sõmera et al., 2021). For this study, we worked with the turnip yellows virus (TuYV) and the sowbane mosaic virus (SoMV), members of two of the four genera regrouped in this family, the *Polerovirus* and *Sobemovirus* genera, respectively.

TuYV has relatively narrow host range but not SoMV (Sõmera et al., 2021). Poleroviruses are transmitted by aphids and *Myzus persicae* has been recently identified as vector for TuYV (Mulot et al., 2018). Sobemoviruses can be transmitted by beetles but transmission by insects belonging to various orders has been reported for SoMV (*Diptera, Hemiptera*) (Sõmera et al., 2021).

A VPg has been found at the 5' extremity of gRNA and sgRNAs in *Solemoviridae* for species of the *Sobemovirus* genus (Hacker and Sivakumaran, 1997; Olspert et al., 2011a, 2011b), but there is no data concerning TuYV and SoMV. The RNA of *Solemoviridae* is not polyadenylated, instead the 3'UTR can fold into stable structures like stem loops as it was reported for the poinsettia latent virus (*Polemovirus*) and the cereal yellow dwarf virus (*Polerovirus*) (Aus Dem Siepen et al., 2005). To date, there is no data on the potential structures found in the 3'UTR of TuYV nor SoMV RNAs.

The genome of poleroviruses encodes seven ORFs. The 5' proximal ORFs code for a suppressor of silencing that functions in triggering AGO1 decay via the autophagy-mediated protein decay (Bellott et al., 2019), for a polyprotein containing domains of a membrane anchor protein, a serine protease, a VPg and a C-terminal protein, as well as the RdRp. The 3' proximal ORFs are expressed from sgRNAs. sgRNA1 translation results in the capsid protein, movement proteins and a factor possibly involved in aphid transmission. The last two ORFs are translated from sgRNA2 and sgRNA3 can also allow the transcription of most 3' proximal ORF (Sõmera et al., 2021) that have been proposed to participate in viral replication and transcription for one *Polerovirus* species (Hwang et al., 2013).

The genome of sobemoviruses presents organizational similarities with poleroviruses. The 5' terminal ORF in the *Sobemovirus* genome also codes for a suppressor of silencing. The nearby ORF encodes for a polyprotein that includes a serine protease, a VPg and a C-terminal protein that has ATPase and RNA helicase characteristics. Ribosomal frameshifting during translation of this ORF allows to produce the RdRp. The last ORF codes for the capsid protein and is translated from a sgRNA (Sõmera et al., 2021).

Tombusviridae

Viral species from this family have a monopartite genome (except species from the *Dianthovirus* genus) contained within a non-enveloped icosahedral viral particle (Rochon et al., 2011). Seventeen

genera are regrouped in this family and for this study, two genera are represented: *Betacarmovirus* with the turnip crinkle virus (TCV) and *Tombusvirus* with the tomato bushy stunt virus (TBSV).

General informations on TBSV and the RNA 3'UTR

Tombusviruses have a narrow host range in nature but under experimental conditions, many plants can by systematically infected such as model plants *N. benthamiana* and *A. thaliana*. Most of the viral species of this genus are soil-borne and can infect a host without the help of a vector. TBSV is a typical member from the *Tombusvirus* genus which genome does not have a cap nor a VPg at its 5' extremity. The non-polyadenylated 3'UTR has been reported to fold into several stem-loop structures in the context of defective interfering RNA (DI RNA) (Fabian et al., 2003; Na et al., 2006) that stimulates the cap-independent translation and replication (Rochon et al., 2011). A more recent study resolved the secondary structure of the whole genome by Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) (Wu et al., 2013a). The 3'UTR contains a Y-shaped 3'CITE followed by three small stem-loops separated by linkers. Both the 3'CITE and the stem-loops are involved in long range interactions with the 5'UTR. The genome ends with a -CCC_{OH}-3' extension that is not involved in any nucleotide interaction within the RNA (Fabian et al., 2003; Na et al., 2006; Wu et al., 2013a). Its genome codes for the replicase protein expressed from the gRNA, a capsid protein translated from a sgRNA and for a movement protein as well as suppressor of silencing, both expressed from another sgRNA (Rochon et al., 2011).

General informations on TCV

The turnip crinkle virus is a betacarmovirus from the *Tombusviridae* family (Rochon et al., 2011). The main symptoms that have first been described upon TCV infection are the crinkled leaves that can be curled at the edge, from which the viral species name originates. In some cases, TCV infection leads to vein clearing and local chlorotic lesions. TCV infects a small host range of species within the *Brassicaceae* family (formerly *Cruciferae*). TCV is transmitted by beetles from the order of *Coleoptera* and can also be inoculated mechanically on many experimental hosts.

Under experimental conditions, TCV induces local lesions or systemic infection in many cruciferous hosts like radish and mustard, as well non-cruciferous hosts like cucumber and tobacco (Broadbent and Heathcote, 1958). TCV also causes local chlorotic lesions in *C. amaranticolor* and *C. quinoa* and can systemically infect model organisms like *Nicotiana benthamiana* and *Arabidopsis thaliana*.

TCV virions are small isometric particles which structure is similar to TBSV viral particle (Hogle et al., 1986). TCV genome replication occurs in viral factories within cellular membrane-derived vesicles, in particular mitochondrial outer membrane (Blake et al., 2007; Russo and Martelli, 1982).

Genomic organization of TCV and associated RNAs

The genome of TCV is composed of a 4 kb RNA segment (Carrington et al., 1989). TCV genome is not capped (Qu and Morris, 2000) and does not have a poly(A) tail nor a TLS. Two subgenomic RNAs of 1.7 and 1.45 kb, respectively, are synthetized during TCV infection (Carrington et al., 1987).

Five overlapping ORFs are encoded by TCV gRNA (Fig. 14C): the first ORF codes for the p28 protein and the read-through translation product p88. Both are involved in the replication of TCV genome

(Hacker et al., 1992). The next two ORFs code for two small proteins, p8 and p9, that are required for viral movement (Hacker et al., 1992; Heaton et al., 1991; Li et al., 1998) and expressed from one of the sgRNAs (Li et al., 1998). The study of recombinant GFP-tagged p8 and p9 proteins showed that p9 is diffuses in the cytoplasm whereas p8 localization is restricted to the nucleus. This is explained by the presence of two nuclear localization signals (NLSs) in p8 sequence (Cohen et al., 2000). The 3' proximal ORF encodes the capsid protein p38 and is expressed from the other sgRNA. The capsid protein is also required for cell-to-cell movement because mutations in this protein results in the abolition of systemic movement (Hacker et al., 1992). TCV capsid protein p38 also suppresses silencing induced by sense, antisense and double-stranded RNA, probably by interfering with DCL activity (Qu et al., 2003; Thomas et al., 2003). P38 also functions in hindering siRNA loading onto AGO1 by directly binding to the protein via GW residues (Azevedo et al., 2010). One study showed that temperature upregulates DCL2 activity and is associated with TCV infection survival in Arabidopsis (Zhang et al., 2012).

Satellite RNAs (satRNAs) have been found to be associated with TCV (Altenbach and Howell, 1981). Such viral elements are either viral particles or viral RNAs that co-infect plants only in the presence of a given virus and for which RNA sequence does not have a high homology with the helper virus. Their replication depends on their helper virus and the host machinery and they can regulate the symptoms (Palukaitis, 2016). In the pioneer study, four satRNAs have been characterized: satRNA A, B, C and D (Altenbach and Howell, 1981). Of note, satRNA C shares the last 3' 166 nt with TCV genomic RNA (Simon and Howell, 1986).

During infection by plant and animal RNA viruses, defective interfering RNAs (DI RNAs) have been detected. These RNAs are truncated versions of viruses, usually lacking some or all ORFs but containing the cis-elements that function in replication of the parental virus (Simon et al., 2004). DI RNAs in tombuviruses have first been described for TBSV (Hillman et al., 1987). The role of DI RNAs remains poorly understood and seems to vary between the viruses (Simon et al., 2004). TCV DI RNAs have been discovered and they have been shown to intensify viral symptoms. Moreover, these DI RNAs result from a *de novo* synthesis in the plant cell (Li et al., 1989).

Recently, a long noncoding RNA (lncRNA) generated during TCV infection in *N. benthamiana* has been reported. It shares sequence similarities with the last 283 nt of TCV genome and was thereby named tiny TCV subgenomic RNA. Both strands of this RNA accumulate in a viral RdRp-dependent manner. The authors identified a new mechanism of viral lncRNA synthesis depending on the polymerase rather than on a cellular exoribonucleolytic decay factor (Zhang et al., 2021).

The 3'UTR organization of TCV genome

This virus has been extensively studied to better understand the role of RNA elements involved in the replication and the translation of viral genomes. The 3'UTR of TCV RNA has been largely characterized and folds into a highly dynamic structure. It contains five hairpins (H4, H4a, H4b, H5 and Pr), three pseudoknots (Ψ1, Ψ2, Ψ3) and a large unstructured region (USR) also called M3H in some papers that begins upstream the stop codon of the last ORF and extend up to the hairpin H4. TCV RNA ends with a 6-nt long tail preceded by the hairpin Pr. A linker of unknown function connects Pr to the hairpin H5 (Fig. 6) (Simon, 2015).

The 3' terminal tail seems to have an important biological function, as it has been highlighted by the different repair mechanisms on TCV satRNA mutants deleted from the tail (Carpenter and Simon, 1996; Guan and Simon, 2000; Nagy et al., 1997). The last four nucleotides of the tail can bind to the large symmetrical loop within H5 and form the Ψ1 pseudoknot. This pseudoknot has been reported to be crucial for the accumulation of satRNA C and TCV gRNA (Zhang et al., 2004, 2006b). Moreover, the abolition of Ψ1 is associated with a higher transcription of the complementary strand, with many transcripts of aberrant sizes (Zhang et al., 2004). These data show that Ψ1 functions as repressor of the minus strand synthesis. It also suggests that this pseudoknot is involved in the correct 3' terminus recognition by the viral RdRp, thereby regulating the viral genome replication. Of note, the authors of this paper also identified similar hairpins in other carmovirus genomes (Zhang et al., 2004).

The GC-rich 3' proximal hairpin Pr is very stable and may be involved in RNA stability. It was first reported in satRNA C to have a core promoter activity in *in vitro* assays (Song and Simon, 1995; Stupina and Simon, 1997). A comparative study between Pr hairpin on both RNAs show that additional elements in the 3'UTR of TCV gRNA regulates the complementary strand synthesis (Zhang et al., 2006b). Mutations in Pr negatively affect transcription and enhance translation (Yuan et al., 2010). In addition, Pr is involved in a long-distance interaction with a bulge loop in a hairpin within the 5' proximal ORF that is essential for stop codon readthrough, generating the RdRp (Cimino et al., 2011). This interaction is conserved in *Tombusviridae* (Simon, 2015).

The H5 hairpin has been mostly studied in satRNA C and it is proposed to be involved in replication. Indeed, it has a large symmetrical internal loop that can inhibits minus strand synthesis by pairing with the four 3' terminal bases of TCV gRNA, thereby blocking the access of the viral RdRp to the 3' terminus. It has been shown that disrupting this interaction stimulates the synthesis of complementary strand *in vitro*. Moreover, mutations in H5 internal loop result in an increased mutation rate (McCormack and Simon, 2004). Accordingly to the work of Zhang and colleagues (Zhang et al., 2004), the association between H5 internal loop and the 3' terminal tail is involved in the regulation of genome replication.

The only 3'UTR elements conserved in all carmovirus genomes are pseudoknot Ψ1 and hairpins Pr and H5 (Simon, 2015). Additional structural elements have been characterized in TCV RNA 3'UTR. H4a and the adjacent upstream sequence can base pair and form the Ψ3 pseudoknot (McCormack et al., 2008; Zhang et al., 2006a), while H4b can pair with a sequence downstream of H5 to form the Ψ2 pseudoknot (Zhang et al., 2006a, 2006c). Both of them are important for viral accumulation *in vivo* (McCormack et al., 2008; Zhang et al., 2006c). Ψ2 is involved in the structural switch to activate replication (Zhang et al., 2006c).

Upstream of the Ψ3 pseudoknot and H4a, there is an A-rich linker sequence that is proposed to be essential for virus accumulation. Indeed, mutations within this linker reduced *in vitro* transcription and translation of a reporter construct (Stupina et al., 2008; Yuan et al., 2009). This linker connects H4a to H4 and has been shown to be important for TCV viral RNA accumulation in protoplasts. H4 hairpin is bound by the viral RdRp and stimulates replication in both orientations (Sun and Simon, 2006).

The unstructured region (USR/M3H) downstream the capsid protein ORF was first described to be involved in recombination and latter for translation enhancement (Carpenter et al., 1995; Stupina et al., 2008). Another interesting role related to NMD resistance has been reported. Deletion of this region

abolished NMD resistance of a reporter gene *in planta*. The authors propose that the NMD resistance is related to lack of secondary structure rather than to the sequence (May et al., 2018).

A T-shaped structure (TSS) formed by the interaction of hairpins H4a, H4b and H5, the A-rich linker as well as pseudoknots Ψ2 and Ψ3 has been described and its structure is similar to a tRNA (Le et al., 2017; McCormack et al., 2008). Despite the fact the secondary structure does not look like the cellular tRNA cloverleaf, the global tertiary structure shares similarities with a tRNA and adopts a twisted T-shape (Le et al., 2017). H5 mimics the anticodon stem and H4a/ Ψ3 mirror the amino acid acceptor stem (Le et al., 2017; Zuo et al., 2010). The H4a/ Ψ3 interaction is very stable (Le et al., 2017). The ribosomes 80S and 60S are able to bind the TSS (Stupina et al., 2008) and mutations that alters the ribosome binding resulted in a decreased translation of a reporter gene, indicating that this structure functions a 3'CITE (Stupina et al., 2008). Moreover, the TSS compete with cellular tRNA for the binding into the P-site of the ribosome and this effect is specific of TCV gRNA (Le et al., 2017; Stupina et al., 2008). H4, the TSS and Pr are key structural elements that regulate the translation and replication of TCV genome (Yuan et al., 2012). Of note, TCV 5'UTR does not have particular structural features and no RNA:RNA interaction between the 5' and the 3'UTR have been observed (Stupina et al., 2011), so a key question lies in the recruitment of the ribosome and the translation initiation. It has been reported that the 40 ribosomal subunit can bind to a poly(UC) tract upstream of p28 start codon. So, the 80S formation by the interaction between the 40S and the 60S subunit leads to TCV RNA circularization and subsequent translation (Stupina et al., 2011).

Tymovirus (Tymoviridae)

The most emblematic species of the *Tymoviridae* family is probably the turnip yellow mosaic virus (TYMV), a tymovirus that we used in this work. Tymoviruses have a narrow host range within the *Brassicaceae* family, but they can infect *A. thaliana*. They are transmitted mechanically or by beetles (Dreher et al., 2009).

TYMV TLS has been extensively studied and can aminoacylated with valine (Dreher, 2010). The structure of TYMV TLS has been resolved by experimental data and prediction. Although the secondary structure does not look like the typical cloverleaf, the 3D structure shares many features with a cellular tRNA (Dreher, 2010; Dumas et al., 1987; Florentz et al., 1982; Rietveld et al., 1983). The first ORF at the 5' end codes for a polyprotein involved in viral replication. Another ORF almost entirely overlapping the first one encodes a protein involved in cell-to-cell movement. The 3' proximal ORF codes for the capsid protein that is translated from a subgenomic RNA (Dreher et al., 2009).

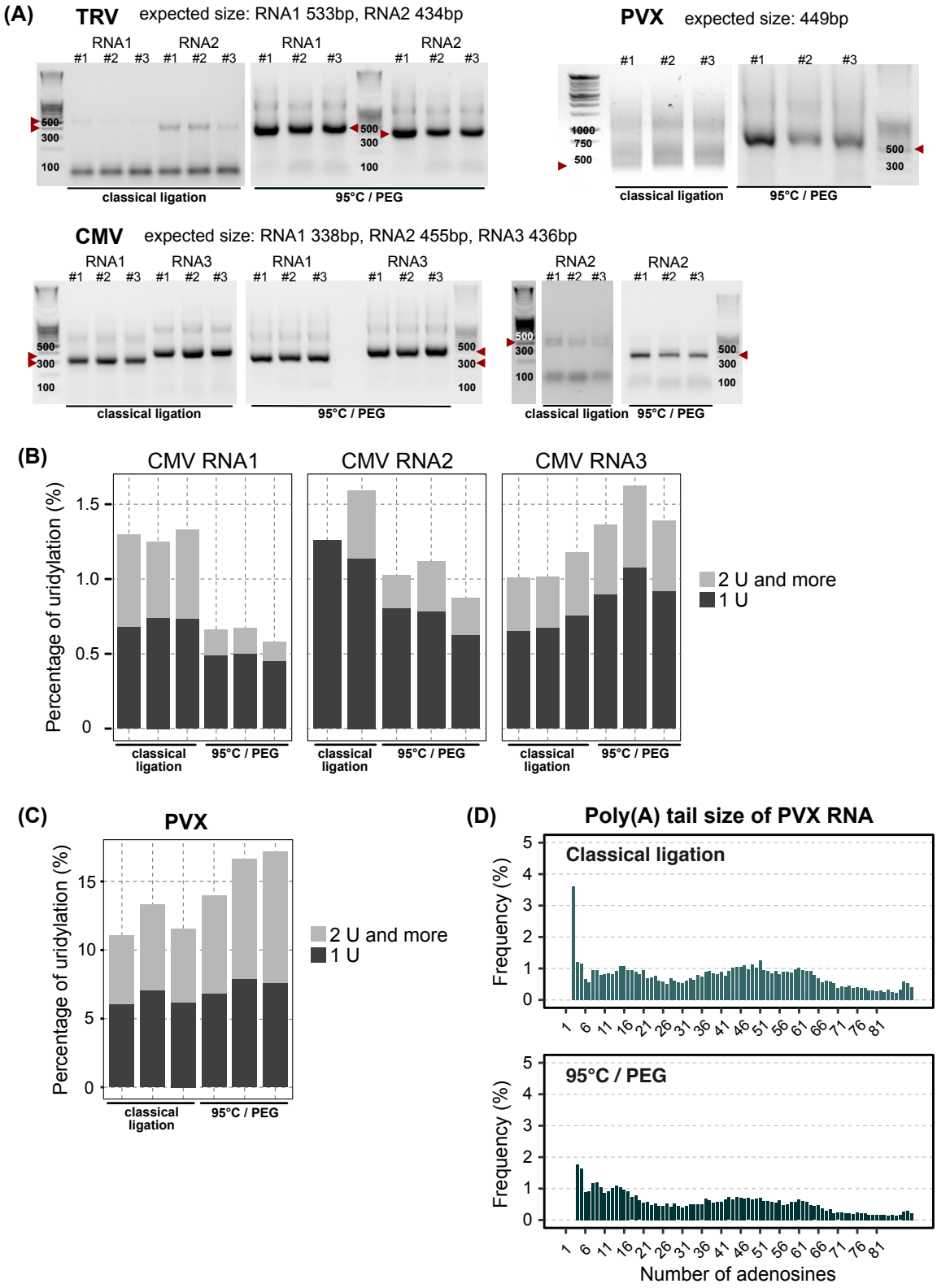


Figure 15: Optimization of the 3'RACE-seq cDNA synthesis for non-polyadenylated RNAs. (A) PCR2 products loaded on a 1.2% agarose gel. Red arrows indicate the product of interest. Sizes of the DNA ladder are written on the corresponding band. **(B) & (C)** Percentage of uridylation of CMV RNAs **(B)** and PVX RNA **(C)** calculated in each plant with the number of reads for each RNA used as denominator. The color code indicates the size of the U-tail: 1 U (dark grey) and 2 Us and more (light grey). Each bar represents one plant. **(D)** Poly(A) tail size distribution of PVX RNA poly(A) tail in one plant as representative example. The frequency is calculated with the total number of reads for PVX RNA in this plant as denominator.

Results

1. A global survey of phytoviral RNA 3' uridylation identifies extreme variations and virus-specific patterns

Uridylation was detected by low-throughput sequencing on viral RNAs from viruses infecting fungi, metazoans and plants (Huo et al., 2016). Recently, one study proposed that uridylation functions in the antiviral defense of animals (Le Pen et al., 2018). Yet, the mechanisms related to viral RNA uridylation remain poorly understood and to date, there is no data on plant viruses. We first wanted to investigate the diversity of plant viral RNA uridylation patterns. To this end, we used a high-throughput sequencing strategy, 3'RACE-seq that allows to determine 3' templated and non-templated nucleotides on target RNAs (Scheer et al., 2020). We selected viral species representative of the main families of viruses with a ss(+)RNA genome infecting plants and in total, we analyzed more than 30 viral RNAs. We also studied the role of cellular TUTases in viral RNA uridylation.

In this chapter, I will first detail an optimization of the 3'RACE-seq protocol to better capture the 3' ends of viral RNAs with a structured 3' region. I will then present a manuscript that recapitulates our results on viral RNA uridylation in plants and show data on other 3' terminal modifications that we detected on plant viral RNAs. Finally, I will discuss the distinct activities of Arabidopsis TUTases on viral RNAs.

1. Optimization of the 3'RACE-seq strategy

3'RACE-seq starts with the ligation of an adapter to the 3' extremity of RNAs with an accessible 3' hydroxyl group using the T4 RNA Ligase 1. The adapter is adenylated at the 5' extremity, so it can be a substrate for the ligase in the absence of ATP (Scheer et al., 2020). 3'RACE-seq is mostly used in our group to study poly(A) tails. The ligation of an adapter can however be prevented if the RNA does not end with a 3' hydroxyl, if there is a particular modification, or if there are strong secondary structures that could hinder the binding of the ligase. When I first started to generate the libraries on viral RNAs, I encountered difficulties to amplify some of them by nested-PCR, even when I have tested different amount of template and alternative PCR conditions. By contrast, the amplification was working when I was using reversed primers that annealed on the viral RNA rather than on the 3' adapter, meaning that amplification troubles were caused by an inefficient adapter ligation. In particular, I encountered such difficulties for some of the viral RNAs that possess strong structures in their 3' UTR, which could prevent the binding of the ligase and/or the accessibility to the 3' hydroxyl group. This was notably the case for TRV RNAs, for which I failed to amplify 3' end regions with our classic 3' RACE-seq protocol (Fig. 15A). To improve the efficiency of the ligation, I denatured the RNAs at higher temperature and I added polyethylene glycol (PEG) 8000 to the reaction. I incubated RNAs 10 s at 95 °C instead of 5 min at 65 °C to improve the denaturation. In addition, I added PEG 8000 at 6.25% (v/v) to the reaction, as PEG 8000 was reported to stimulate the ligation reaction by a ten-fold rate (Harrison and Zimmerman, 1984).

This tip was given by Dr Anthony Gobert (IBMP), and it allowed to obtain 3' RACE-seq libraries for TRV RNAs (Fig. 15A).

To compare if the information on the 3' non-templated nucleotides is similar with the two protocols, I also performed 3' RACE-seq libraries using the two strategies of ligation for two other viruses: a polyadenylated RNA virus, PVX, and a non-polyadenylated RNA virus, CMV (RNA1-3). I first compared the efficiency of amplification by looking at PCR profiles. Whereas no difference was detected for CMV RNA1 and RNA3, the amplification for PVX RNA and CMV RNA2 was more efficient with the optimized ligation protocol (high denaturation temperature + PEG addition) (Fig. 15A). These results suggest that, at least for some viral RNAs, a stronger RNA denaturation and the addition of PEG allow to improve the ligation and thus, cDNA synthesis and subsequent PCR amplification efficiency.

Next, we ought to determine whether the different strategies of ligation resulted in a difference in the detection of non-templated nucleotides, in particular uridines. For CMV RNAs, the measured level of terminal uridylation is between 1 and 1.5% with the classical ligation protocol (Fig. 15B). Of note, there were not enough sequences of CMV RNA2 to plot the percentage of uridylation for the third plant. For the samples generated with the optimized ligation protocol, we detect terminal uridylation on CMV RNA2 at a similar level in the three plants, as well as for RNA1 and RNA3 (Fig. 15B). We observe that the percentage of uridylation decreased to 0.5% for CMV RNA1 when the optimized protocol is used. This decrease could mean that we captured more CMV RNA1 molecules that are not modified at the 3' extremity. Since this protocol improves the ligation efficiency, capturing more non-tailed molecules could mean that CMV RNA1, if not tailed by non-templated nucleotides, is hardly accessible to the ligase, because of strong structures for example. For PVX RNA, we detected slightly more uridylated reads for the samples generated with the optimized ligation protocol (Fig. 15C). These differences are still minor and could be technical.

To determine whether the optimized ligation strategy impact the poly(A) tails, I also tested this protocol with PVX RNA which is polyadenylated. The distribution of the poly(A) tail of PVX RNA was plotted for poly(A) tails ranging from 3 to 90 As (Fig. 15D). I removed the tails of 1-2 As because their high number masks the distribution of the rest of the sizes. Tails longer than 90 nucleotides are removed as well because it corresponds to the maximum number of adenosines measured in the poly(A) tail sequence. Indeed, among the 111 nucleotides sequenced in the read 2, used to analyze RNA 3' end and non-templated nucleotides, 20 are from the 3' adapter and all tails longer than 90 As have a measured size of 91 As. Looking at the poly(A) tail size distribution, we see that PVX RNA poly(A) tail ranges from 3 to 70 As, with two peaks at 12 and 50 As, respectively. We do not see substantial differences between the poly(A) tail profiles between the two ligation strategies (Fig. 15D).

These first results comparing two ligation strategies showed that the optimized strategy improves the PCR efficiency, probably a consequence of an improvement of the ligation efficiency. We did not see major differences in the detection of 3' non-templated nucleotides, whether it is for a polyadenylated RNA, PVX RNA, or for non-polyadenylated RNAs, CMV RNAs, even if there is slightly less variability for CMV RNA2. This could be explained by the improvement of the PCR efficiency, resulting in the capture of more CMV RNA2 molecules. So, we decided to homogenize library preparation for the non-

polyadenylated RNAs by performing all the ligations with the PEG strategy, while the polyadenylated RNAs were ligated to the adapter using the classical 3'RACE ligation protocol.

2. Article

a. Personal contributions

The data presented in the above manuscript have been obtained in collaboration with the team of Olivier Lemaire (SVQV, INRAE Colmar, France). I was substantially involved in the organization of the experiments, the generation of infected material and the collection of samples.

For the analysis of the uridylation landscape of plant viral RNAs, I generated 3'RACE-seq libraries for non-polyadenylated viral RNAs and for some of the polyadenylated viral RNAs. Shahinez Garcia (SVQV, INRAE Colmar, France) generated libraries for viral RNAs from species of the *Secoviridae* and some of the grapevine-infecting viruses. We worked closely together during libraries preparation. To assess the role of cellular TUTases in viral RNA uridylation, I performed viral inoculations, sample collection and libraries preparation. I did the data analyses under the supervision of Hélène Zuber and she helped me to fine-tune the data analysis.

b. Main text: A global survey of phytoviral RNA 3' uridylation identifies extreme variations and virus-specific patterns

Anne Caroline Joly[1], Shahinez Garcia[2], Jean-Michel Hily[2], Sandrine Koechler[1], Damien Garcia[1], Emmanuelle Vigne[2], Olivier Lemaire[2], Hélène Zuber[1,*] and Dominique Gagliardi[1,*]

[1]Institut de biologie moléculaire des plantes, CNRS, Université de Strasbourg, 67000 Strasbourg, France.

[2]Santé de la Vigne et Qualité du Vin, INRAE, Université de Strasbourg, 68000 Colmar, France.

[*]Correspondence:

dominique.gagliardi@ibmp-cnrs.unistra.fr

+33367155366

helene.zuber@ibmp-cnrs.unistra.fr

+33367155381

Keywords

Uridylation, TUTase, RNA degradation, deadenylation, GFLV, TuMV, grapevine

Abstract

Viral RNAs can be uridylated across eucaryotic hosts, including plants. Yet, our knowledge on uridylation patterns and roles remains rudimentary for phytoviruses. Here, we report global 3' terminal RNA uridylation profiles for representatives of the main families of positive single-stranded RNA phytoviruses. Uridylation patterns are remarkably variable even for viruses within the same genus. Uridylation marks degradation intermediates, but also full-length viral RNAs. Surprisingly, most poly(A) tails of encapsidated *Grapevine fanleaf virus* (GFLV) RNAs are strictly mono-uridylated. This peculiar mono-

uridylation of poly(A) tails is independent of host TUTases and is a hitherto unknown type of viral genomic RNA extremities. Mono-uridylated poly(A) tails are not widespread among nepoviruses, but restricted to GFLV and its closely related *Arabis mosaic virus* (ArMV), both linked to grapevine fanleaf degenerative disease. Our work unveils the extreme diversity of uridylation patterns across phytoviruses and constitute a valuable resource to decipher pro- and anti-viral roles of uridylation.

Introduction

Viruses represent a constant threat to human health and food security worldwide. The development of effective antiviral strategies relies on understanding the molecular processes associated with the viral cycle and host-pathogen interactions. In plants, three main lines of defense have evolved to fight viral infections: physical barriers that viruses must overcome to penetrate into cells, the innate immune response and RNA silencing[1]. The RNA degradation machinery can also interfere with viral infections by adjusting the transcriptome and directly targeting viral RNAs. Recently, RNA uridylation was proposed as an antiviral defense mechanism in animals[2]. RNA uridylation corresponds to the addition of one to several uridines at the 3' end of an RNA. This reaction is catalyzed by terminal uridylyltransferase (TUTases) and this post-transcriptional process is conserved across eucaryotes, except *Saccharomyces cerevisiae*[3-7]. Uridylation targets both non-coding RNAs and mRNAs, and its primordial role is to induce RNA degradation[3-7]. Interestingly, a genetic screen identified the TUTase CDE-1 as a resistance factor for the *Orsay virus* (OrV) in *Caenorhabditis elegans*[2]. CDE-1 was proposed to uridylate OrV RNA to facilitate its degradation. Similarly, TUT4 and TUT7, two cytosolic TUTases in the cytosol of human cells, repress the expression of influenza A virus mRNA and protein levels[2]. Altogether, these observations led to the conclusion that RNA uridylation acts as an antiviral defense mechanism in animal cells[2].

The uridylation of viral RNAs is unlikely restricted to animals and in fact, long heteropolymeric U-rich nucleotide extensions were detected for viruses in several eucaryotic species including plants[8]. However, this detection of viral RNA tailing was biased by the use of a low-throughput sequencing analysis of clones obtained by priming cDNA synthesis with an oligo(dA) primer. To obtain a global view of phytoviral RNA tailing, we used a high-throughput sequencing strategy to survey 3' nucleotide addition to viral RNAs from representatives of the main families of single-stranded positive (ss(+)) RNA viruses infecting plants. Our results reveal an unexpected diversity in phytoviral RNA uridylation patterns. We show that different activities can uridylate viral RNA degradation intermediates or genomic RNAs and, in light of these results, we discuss the potential pro- and anti-viral roles of RNA uridylation.

Results and discussion

High resolution mapping of phytoviral RNA 3' ends by 3'RACE-seq

To explore the diversity of phytoviral RNA tailing, we initially selected 21 viruses representing 7 of the 8 orders of ss(+) RNA phytoviruses (Table 1). These representative viral RNAs cover the main types of known phytoviral 3' extremities such as aminoacylatable and non-aminoacylatable tRNA-like structures (TLS), various non-TLS 3' terminal structures and poly(A) tails. Because poly(A) tails can be

either encoded or extended by poly(A) polymerases, we refer hereafter to 3' tailing as for the addition of nucleotides 3' to the poly(A) tail itself, irrespective of the mode of synthesis of the poly(A) tail.

Some of these 21 viruses have multipartite genomes and overall, 31 viral RNAs were analyzed by 3'RACE-seq (Table 1). This Illumina-based protocol allows for in-depth mapping of RNA 3' extremities with single nucleotide resolution, including the identification of any untemplated nucleotides. Of note, Unique Molecular Identifiers (UMIs) are incorporated by ligating the 3' ends of RNAs to an adapter that contains a 15-nucleotide degenerate sequence. This molecular barcoding allows for a deduplication step during data analysis and therefore each final read corresponds to a single original RNA molecule. 3'RACE-seq has a few technical limitations that are worth to note in the context of this study. The first one is that 3'RACE-seq interrogates 3' extremities independently of 5' extremities and therefore, full-length viral RNAs and subgenomic RNAs sharing identical 3' extremities are not discriminated here. Also, the 3' ends of the RNA targets must be accessible to the T4 RNA ligase used to ligate the adapter. Therefore, certain modifications of the last nucleotide, like aminoacylation of tRNA-like structures (TLS), will prevent ligation. Yet, the fraction of non-aminoacylated viral RNAs can still be analyzed. Finally, 3' terminal secondary structures may also impede T4 RNA ligase. To limit this effect, non-polyadenylated viral RNAs were denatured at 95°C before the ligation step. Despite these few constraints, 3'RACE-seq remains a powerful method to accurately map 3' extremities of target RNAs.

The expected 3' ends and polyadenylated/non-polyadenylated statuses were confirmed for 19 out of the 21 viruses. However, the 3' terminal features of both the *Carrot necrotic dieback virus* (CNDV, *Secoviridae* family, *Sequivirus* genus) and the *Grapevine leaf roll-associated virus 2* (GLRaV-2, *Closteroviridae* family, *Ampelovirus* genus) RNAs were reassessed. The CNDV RNA was proposed to be polyadenylated because its 3' region could be amplified by RT-PCR using oligo(dT)-primed cDNA[9]. It is possible that such an amplification was due to a minor fraction of CNDV RNA being polyadenylated or because this viral RNA ends with an A-rich region. However, our 3'RACE-seq data demonstrate that its 3' extremity is not constitutively polyadenylated. Indeed, 99.8% of the reads that map to the last nucleotide of the CNDV reference sequence are not tailed (Fig. 1a, 1b). The CNDV RNA will therefore be classified amongst the non-polyadenylated viral RNAs hereafter.

The current description of the 3' terminal features of the GLRaV-2 RNA is rather contradictory. The GLRaV-2 RNA is described as not polyadenylated by several resources gathering general information on viruses, like the International Committee on Taxonomy of Viruses (ICTV) or ViralZone[10]. Yet, a BLAST search retrieved several GLRaV-2 sequences ending with either a short poly(A) tail or a longer A-rich region followed by a short nucleotides stretch like GAAGC or GCGGCCGC[11,12]. Our 3'RACE-seq experiment revealed that most (90.4%) of GLRaV-2 RNA 3' ends correspond to A-rich tails (Fig. 1c-1e). Only a minority (8.7%) of these tails are homopolymeric which is in contrast to the pure adenosine extensions of 8 other selected viruses whose RNAs are polyadenylated. The GLRaV-2 RNA A-rich tails often contains Us (36.2% of the tails) and also frequently Gs and Cs in the last 3' nucleotides (57.5% of the tails). A sequence logo analysis indicates that those tails often terminate with GAAGC, as previously reported for a GLRaV-2 infectious clone[12] (Fig. 1f). Overall, our 3'RACE-seq analysis reveals that GLRaV-2 RNAs have complex 3' A-rich tails. It is likely that homopolymeric poly(A) tails and A-rich tails of phytoviral RNAs are produced by different enzymes or distinct mechanisms yet to be elucidated.

However, it is unknown at present whether they would entail different functions. In any case, GLRaV-2 RNA is considered hereafter in a specific sub-class of polyadenylated viral RNAs with A-rich tails.

Extreme diversity of RNA uridylation levels across ss(+) RNA phytoviruses

The diversity of 3' terminal nucleotide addition was then analyzed for the 21 selected viruses. Three main observations were made. Firstly, 3' tailing of plant viral RNAs is widespread. Overall, viral RNAs can be adenylated, cytidylated, guanylated, uridylated or tailed with mixed nucleotide extensions (Fig. 2a, Supplementary Fig. 1 and Supplementary Data 3b). The second general observation is that an extreme variability in uridylation levels was detected for viral RNAs, ranging from 0.2% to 90% (Fig. 2a). Of note, we cannot exclude that the smallest level of uridylation detected for CNDV RNA may correspond to background level, with no biological significance. Also, these percentages refer to uridylation detected by 3'RACE-seq in the 3' most terminal region of viral RNAs and additional upstream uridylation sites may also exist but are not considered here. The third and last general observation is that 3' terminal uridylation patterns are varied, from strict mono-uridylation for some viral RNAs to tails up to 30 uridines, which is the maximal size measured with our analysis pipeline (Fig. 2b). This striking variability in uridylation levels and patterns likely indicates distinct roles in viral RNA metabolism. We therefore decided to focus our study on further analyzing phytoviral RNA uridylation.

Phytoviral RNAs with a TLS are poorly uridylated

We first analyzed the uridylation patterns of 18 non-polyadenylated viral RNAs representing 12 phytoviruses (see Table 1 for genomic features). Those phytoviral RNAs are uridylated from 0.2% up to >10% (Fig. 2a). Seven of the 12 selected phytoviruses have a relatively low level of uridylation (<2.2%): the *Peanut clump virus* (PCV), the *Cucumber mosaic virus* (CMV), the *Turnip yellows virus* (TuYV), the *Tobacco mosaic virus* (TMV), the *Alfalfa mosaic virus* (AlMV), the *Turnip yellow mosaic virus* (TYMV) and the CNDV (see classification in Table 1). Interestingly, four of these seven viruses have a known 3' terminal TLS: PCV, CMV, TMV and TYMV (Fig. 2a). TMV TLS can be aminoacylated with histidine, PCV and TYMV TLS with valine, and CMV TLS with tyrosine[13]. Aminoacylation will prevent uridylation, and the extent of the competition between uridylation and aminoacylation is difficult to estimate *in planta*. Yet, among those viral RNAs that remain unaminoacylated *in vivo* and therefore detectable by 3'RACE-seq, only a minor proportion is uridylated (from 0.3 to 2.2 %) (Fig. 2a). Hence, those TLS seem to be poor substrates or poorly accessible to the TUTase(s) responsible for uridylating viral RNAs.

The three other viral RNAs with a low level of uridylation are the TuYV, the CNDV and the AlMV RNAs. To our knowledge, the structure of the 3' terminal region has not been determined for TuYV and CNDV RNAs. By contrast, the 3' terminal region of AlMV is known to switch between two alternative conformations[14]. Either five 3' terminal stem-loop structures bind the coat protein favoring translation, or a pseudoknot allows a conformational rearrangement to generate a structure resembling a TLS and this pseudoknot is necessary for replication[14]. This TLS-like conformation, but also probably the binding of the coat protein, may restrict access to TUTases, thereby explaining the low uridylation rate of the three AlMV RNAs (Fig. 2a).

Uridylation of degradation intermediates reveals patterns of ribonucleolytic attacks

The five other phytoviruses with non-polyadenylated RNAs have a relatively high level of uridylation (from 3.4 to 12 %) (Fig. 2a): the *Grapevine leaf roll-associated virus 1* (GLRaV-1), the *Turnip crinkle virus* (TCV), the *Sowbane mosaic virus* (SoMV), the *Tomato bushy stunt virus* (TBSV), and the *Tobacco rattle virus* (TRV) (see classification in Table 1). SoMV and TRV RNAs are uridylated mostly at their mature extremities, similarly to what is observed for the TLS-ending RNAs of PCV, CMV, TMV and TYMV, and for the three AlMV RNAs (Fig. 3a, Supplementary Fig. 2). Therefore, the higher uridylation rates for SoMV and TRV RNAs as compared to viral RNAs ending with a TLS likely reflect a greater accessibility of their 3' extremities by TUTase(s). Yet, the accessibility of the 3' terminal extremities is not the only feature regulating uridylation as we also detected uridylation sites located upstream of mature 3' extremities. Certain of these internal uridylation sites are not conserved between replicates as for PCV RNA1 and RNA2, or are scarce as for SoMV and TuYV RNAs (Fig. 3a, Supplementary Fig. 2). By contrast, robust uridylation patterns of truncated viral RNAs that are well conserved across biological replicates were identified for TBSV, TCV and GLRaV-1 RNAs (Fig. 3a). Because of the internal position of these uridylation sites and the current knowledge of the primordial role of uridylation in triggering RNA degradation in the cytoplasm of eukaryotes, including viral RNAs in *C. elegans*[2], we propose that these truncated uridylated viral RNAs represent degradation intermediates.

The GLRaV-1 RNA displays some ragged 3' extremities in a ca 100 nt window upstream of its major 3' extremity and those extremities correspond to the main uridylation sites (Fig. 3a). It is therefore possible that GLRaV-1 RNA 3' extremities are subjected to repeated cycles of uridylation and exoribonucleolytic nibbling to overcome stabilizing features like stably bound proteins or structural elements. To our knowledge, such features are not yet characterized for GLRaV-1. By contrast, the structure of TCV RNA 3' region has been studied intensively. TCV RNA 3' UTR begins with a region of ca 49 nt that was first termed the unstructured region (USR) albeit it was later shown to contain a weakly structured hairpin, named M3H[15,16]. Immediately downstream of USR/M3H are five stable hairpins (H4, H4a, H4b, H5 and Pr) and three H-type pseudoknots (Ψ1, Ψ2, Ψ3)[16,17]. The pseudoknots H4a/ Ψ3 and H4b/ Ψ2 together with H5 fold into a T-shaped structure (TSS)[18] (Fig. 3b). The strong structure of TCV RNA 3' region clearly influences its uridylation pattern: three highly reproducible clusters of uridylation sites are detected within the last 200 nt (Fig. 3b). The first cluster corresponds to the USR/M3H region which is positioned exactly upstream of the H4 hairpin. The second cluster corresponds to the loop of H4b hairpin and the third and last cluster to the 3' most terminal 5 nucleotides (UGCCC) of TCV that are immediately downstream of the G/C-rich Pr hairpin. Those terminal nucleotides are likely either directly accessible to TUTase(s) or generated by nibbling of mature 3' extremities up to the Pr hairpin. Very few uridylation sites are detected between the three clusters suggesting that hardly any degradation intermediates are generated in those regions. Rather, the clusters of uridylation sites upstream of H4 and in the H4b loop indicate that these sites could correspond to 3' extremities of degradation intermediates, which may be generated by host endoribonuclease(s).

Because of their robustness across biological replicates, the uridylation patterns of TBSV, TCV and GLRaV-1 RNA degradation intermediates represent signatures of either endoribonucleolytic or 3'-5' exoribonucleolytic attacks that generate truncated RNAs. Those signatures contribute to understanding

3'-5' degradation processes of phytoviral RNAs, especially in light of structural data of their 3' region. Of note, the three uridylation clusters detected when TCV infects *Nicotiana benthamiana* are also detected in Arabidopsis (Supplementary Fig. 3). However, the positions of the main uridylation sites are shifted in the clusters I and III (Supplementary Fig. 3). Hence, the degradation signatures of TCV RNA vary between the two host plant species, likely reflecting differences in RNA degrading activities. For instance, the Arabidopsis Col-0 genetic background used in this study lacks an active SOV[26], the plant ortholog of DIS3L2, which preferentially degrades uridylated RNAs[27,28].

Host TUTases differentially uridylate TCV RNA

The characteristic uridylation patterns of TCV RNA make this viral RNA an adequate model to test whether and which TUTases of the host plant are responsible for its uridylation. Two TUTases have been characterized in *Arabidopsis thaliana*, URT1 and HESO1[19,20]. Although both enzymes may cooperate in uridylating common RNA substrates, they have marked preferences: URT1 is the main TUTase uridylating mRNAs, whereas HESO1's main substrates are small RNAs and RISC-cleaved mRNAs[19-25]. HESO1 was proposed to synthesize longer tails than URT1[22]. Longer tails may be U-rich rather than homopolymeric, because TUTases may infrequently incorporate A, G or C. Because U-rich tails were detected on TCV RNA (Supplementary Fig. 1c), we compared the profiles for both only-U and U-rich tails after TCV infection of wild-type (Col-0) plants, the *urt1-1* and *heso1-4* single mutants, as well as the double mutant *heso1-4 urt1-1*. The only-U and U-rich tails were detected at similar positions in wild-type plants, allowing us to define three independent uridylation clusters in the 3' region of TCV RNA (Fig 4b,c).

Interestingly, while only-U tails slightly decrease in the *heso1-4* mutant, U-rich tails drastically drop, demonstrating that HESO1 uridylates TCV RNA and has a predominant role in the addition of U-rich tails (Fig. 4a-e). TCV RNA uridylation is abrogated in the double mutant *heso1-4 urt1-1*, indicating that URT1 can also uridylate TCV RNA. Yet, URT1 and HESO1 are not fully redundant because distinct 3' extremities were uridylated in the respective *urt1-1* and *heso1-4* single mutants (Fig. 4b). For instance, the 3' terminal uridylation (cluster III) is unchanged in the *urt1* mutant as compared to wild type, but almost abrogated in the *heso1-4* mutant (Fig. 4b,e). Therefore, the 3' terminal uridylation of TCV RNA is mostly catalyzed by HESO1. Conversely, URT1 uridylates preferentially some positions in clusters I and II (Fig. 4b-d). In most cases, URT1 uridylates TCV RNA 3' extremities terminating by As, whether HESO1 seems to prefer 3' extremities ending with non-A nucleotides (Fig. 4f). This differential uridylation of TCV RNA by both TUTases actually reflects their known *in vitro* and *in vivo* substrate specificities[19,22].

By using TCV RNA as model, our data identify a viral RNA as another substrate shared by both URT1 and HESO1, albeit with different preferences for uridylation sites and tail composition.

Strict mono-uridylation to mRNA-like uridylation patterns of polyadenylated phytoviral RNAs

One of the most striking findings from our global survey of phytoviral RNA tailing was the unexpected amplitude of uridylation levels among phytoviral RNAs that are polyadenylated (Fig. 2a). Those viruses correspond to the *Grapevine fanleaf virus* (GFLV), the *Potato virus X* (PVX), the *Grapevine redglobe*

virus (GRGV), the *Grapevine rupestris stem pitting-associated virus* (GRSPaV), the *Grapevine pinot gris virus* (GPGV), the *Grapevine virus B* (GVB), the *Turnip mosaic virus* (TuMV), the *Beet necrotic yellow vein virus* (BNYVV) and the *Grapevine leafroll-associated virus* 2 (GLRaV-2) (see classification in Table 1).

GFLV RNA uridylation patterns have three remarkable features as compared to all other polyadenylated viral RNAs investigated here. Firstly, GFLV RNAs are uridylated at >81% (Fig. 2a). We exclude that the high uridylation level of both GFLV RNAs is due to replication in grapevine because similar uridylation levels were detected when Arabidopsis was used as host for GFLV (compare Figures 2 and 7). Secondly, GFLV RNAs are strictly mono-uridylated (Fig. 2b), and thirdly, uridylated and non-uridylated GFLV poly(A) tails have similar sizes (Fig. 5). To our knowledge, no endogenous RNA has a similar uridylation pattern in plants.

As compared to GFLV RNAs, all other polyadenylated viral RNAs investigated here are uridylated to a much lower extent (Fig. 2a) and those RNAs can be uridylated by one to several uridines (Fig. 2b). In addition, the uridylated poly(A) tails are significantly shorter as compared to non-uridylated ones (Fig. 5). This reduction in size is particularly obvious for TuMV RNA. Unexpectedly, the size of oligo(A) tails that are uridylated is variable across viral RNAs. For instance, uridylated oligo(A) tails are larger for PVX than for TuMV (Fig. 5). These virus-specific uridylation patterns suggest a more complex interplay between deadenylation and uridylation processes than was previously observed for mRNAs in Arabidopsis[19,24,25].

Distinct involvement of host TUTases in uridylating TuMV RNA

TuMV was chosen to evaluate the respective involvement of URT1 and HESO1 in uridylating a polyadenylated phytoviral RNA. Uridylation levels significantly drop in a *urt1-1* mutant but not in a *heso1-4* mutant, indicating that URT1 has a predominant role in uridylating TuMV RNA (Fig. 6a). Yet, uridylation is almost abrogated in the double mutant *heso1-4 urt1-1*, revealing a secondary role for HESO1. Interestingly, the oligo(A) tails uridylated by HESO1 are shorter than for URT1 (Fig. 6b,c). The median sizes of oligo(A) tails uridylated in wild type and in the *heso1-4* mutant are 10 and 11 nt, respectively, whereas this size drops to 4 nt when only HESO1 uridylates TuMV oligo(A) tails in the *urt1-1* mutant. This difference in uridylated oligo(A) tail size could reveal a sequential action of both TUTases. Also, the respective size of the oligo(A) and the U-tails could influence what factors bind to the corresponding TuMV RNA, and therefore URT1- or HESO1-mediated uridylation could trigger different fates for TuMV RNA.

GFLV RNAs are not uridylated by known host TUTases

By contrast to TuMV RNA, the atypical uridylation patterns of GFLV RNAs did not support the involvement of neither URT1 nor HESO1. Indeed, the high uridylation level of both GFLV RNAs is unaffected by the lack of either URT1, HESO1 or both TUTases (Fig. 7a, Supplementary Fig. 4). To determine whether the single 3' terminal uridine of the positive strand is encoded and whether this feature is shared by the negative strand, we attempted to map the 5' and 3' ends of the minus strand

for both GFLV RNA1 and RNA2. The 3' end of GFLV RNA negative strands was analyzed by 3'RACE-seq. For both RNA1 and RNA2 negative strands, most reads (70.1%) map to the expected 3' terminal nucleotide, which is a uridine (Fig. 7b, Supplementary Fig. 5a). This 3' terminal uridine of both minus strands is complementary to the previously mapped 5' terminal adenosine of each positive strand[29,30]. To check whether the poly(A) tail and 3' terminal uridine of the positive strands are also encoded, we then mapped the 5' extremity of each negative strand. Those extremities are not accessible to ligation (data not shown), presumably because of the presence of a viral protein genome-linked (VPg). Therefore, a complementary DNA strand was synthesized to each RNA negative strand, and the cDNA 3' extremities were mapped by an adapted RACE-seq protocol (Supplementary Fig. 6). The 5' extremities of RNA1 and RNA2 positive strands were simultaneously mapped to validate the method. The RACE-seq results confirmed the 5' extremity of both positive strands as previously reported[29,30] (Supplementary Fig. 5b). Interestingly, oligo(U) of up to 50 Us were detected as 5' sequences of both RNA1 and RNA2 negative strands (Fig. 7b, Supplementary Fig. 5c). However, we never observed a 5' terminal adenosine to the oligo(U) for the negative strands which would have indicated that the 3' terminal uridine of the positive strands is be encoded.

Overall, our data show that uridylation of the viral genome is independent of the host TUTases URT1 and HESO1 and that both negative and positive GFLV RNA1 and RNA2 strands terminate with a 3' uridine. In addition, our results indicated that the 3' terminal uridine of the negative strands is encoded by the 5' terminal adenosine of the positive strands and that at least part of the poly(A) tail of the positive strands is encoded by oligo(U) sequences of the negative strands. Even if a negative result cannot be interpreted as proof, we could not find evidence that the 3' terminal uridine of the positive strands is encoded by the minus strand. Either this uridine is added by a yet unknown terminal nucleotidyl transferase activity of the plant host or this mono-uridylation is performed by a viral factor. In fact, the addition of untemplated nucleotides has been reported for several viral RNA polymerases from various viruses including picornavirus, calicivirus, flavivirus, nodavirus, alphavirus, hepacivirus, vesiculovirus, coronavirus and bacteriophage ϕ6[31-42]. Thus, the GFLV RNA polymerase is among the candidate factors to test whether it could add a single uridine after synthesizing the poly(A) tail of the positive strands.

Unique high uridylation rates of GFLV and ArMV among Secoviridae

Two members of the *Secoviridae* family, CNDV and GFLV, were among the initial selection of viruses analyzed in Fig. 2. Yet, those viral RNAs have very distinct features. The single CNDV RNA is not polyadenylated and hardly uridylated, whereas both GFLV RNAs are polyadenylated and uridylated to high levels. To evaluate the evolutionary conservation of GFLV 3' terminal features, we selected 8 other representatives of the *Secoviridae* family. Those viruses are the *Cowpea mosaic virus* (CPMV), the *Broad bean wilt virus 1* (BBWV-1), the *Raspberry ringspot virus* (RrRSV), the *Tobacco ringspot virus* (TRSV), the *Arabis mosaic virus* (ArMV), the *Tomato black ring virus* (TBRV), the *Cherry leaf roll virus* (CLRV) and the *Strawberry latent ringspot virus* (SLRSV) (see Table 2 for classification and RNA 5' and 3' terminal features). Except for CNDV, all other *Secoviridae* including GFLV have two genomic RNAs, that are described as polyadenylated. The evolutionary relationship between these viruses is illustrated by a phylogenetic tree built using the aminoacid sequence of the conserved protease-polymerase (Pro-

Pol) region used by ICTV to define *Secoviridae* species (Fig. 8, Supplementary Fig. 7). For technical reasons this experiment was mostly performed using quinoa and/or grapevine as hosts (Fig 8). Although this latest analysis of GFLV RNA uridylation by 3'RACE-seq confirmed their high uridylation rates as compared to other viral RNAs, an intermediate level of uridylation was observed for GFLV RNA1 and RNA2 especially in the third replicate of infected quinoa plants. This decrease in uridylation is not caused by a lower RNA quality of this sample, as we did not observe any correlation between RNA quality and lower uridylation percentages (data not shown). This observation raises the interesting possibility that GFLV RNA uridylation levels might be regulated by a yet unidentified condition *in planta*. Interestingly, only GFLV and ArMV RNAs share a high level of mono-uridylation among *Secoviridae* (Fig. 8). Even RpRSV and TRSV are uridylated to very low levels (<0.8%) as compared to GFLV and ArMV, despite all these viruses belong to the same sub-group A of the genus *Nepovirus* (Table 2, Fig. 8 and Supplementary Fig. 7). Finally, we determined the uridylation status of GFLV genomic RNAs, by analyzing GFLV RNA1 and RNA2 extracted from purified virions. This analysis revealed that uridylation can reach up to 96.9% of the encapsidated GFLV RNAs (Fig. 9). Therefore, mono-uridylation is a new genomic RNA feature among phytoviral RNAs, that is shared by the two closely related nepoviruses GFLV and ArMV. Both viruses are the causal agents of grapevine fanleaf degenerative disease, a yet uncurable disease causing massive yield losses in the wine industry worldwide[43-45].

Concluding remarks

We report here that RNA uridylation is a widespread modification across phytoviruses. Unexpectedly, phytoviral RNAs show extreme variations in uridylation patterns. A peculiar observation is the high frequency of mono-uridylation for the nepoviruses GFLV and ArMV, that potentially has rather pro-, than anti-viral functions. This hypothesis is supported by the high uridylation rate of encapsidated GFLV RNAs. In addition, the mono-uridylation of GFLV RNAs is not mediated by the two host TUTases reported to facilitate RNA decay. The addition of a single uridine by the viral RNA polymerase after completion of the poly(A) synthesis (which is at least partly templated by an oligo(U) sequence) is among the possible scenarios. Whether the mono-uridylation of GFLV and ArMV is linked to viral replication and/or is relevant for their infectivity needs to be investigated. By contrast, for most other phytoviral RNAs tested, uridylation was frequently detected on degradation intermediates, such as truncated or oligo-adenylated viral RNAs. For those viruses, uridylation may stimulate phytoviral RNA degradation as it does for coding and non-coding cellular RNAs, and as it was proposed for the Orsay virus in *C. elegans* and the influenza A virus mRNAs in human cells[2]. The molecular roles of uridylation during viral RNA degradation seem complex in plants, as we have identified distinct patterns of uridylation for the TUTases URT1 and HESO1, both for TCV and TuMV RNAs. Because URT1 and HESO1 add different lengths of U-extensions and connect distinct cellular factors, uridylation may favor viral RNA degradation via cooperative pathways that need to be further explored. Finally, we recently reported that URT1-mediated uridylation prevents the excessive deadenylation of Arabidopsis mRNAs, which otherwise favors spurious siRNA biogenesis targeting endogenous mRNAs[25]. Therefore, the uridylation of oligo-adenylated viral RNAs as observed for TuMV may play an analogous role: in a wild-type plant, preventing excessive deadenylation could assist TuMV RNA to escape detection by the silencing

machinery. Studying viral RNA uridylation at each step of the virus cycle would be essential to decipher all pro-and anti-viral roles of uridylation.

Methods

Plant growth conditions. The *Arabidopsis thaliana* plants used in this work are of Columbia accession (Col-0). Arabidopsis AGI analyzed in this study were AT2G45620 (*URT1*) and AT2G39740 (*HESO1*). Arabidopsis mutants analyzed in this study are T-DNA insertion lines: *urt1-1* (SALK_087647C) [19] and *heso1-4* (GK-369H06-017072). *heso1-4 urt1-1* was obtained by crossing and provided by P. Brodersen's lab (University of Copenhagen, Denmark). For inoculation experiments, plants are grown on soil with 12 h light / 12 h darkness cycles in a neon-lit chamber (at 21/18 °C). *Nicotiana benthamiana* (wild-type and the *35S::B2:*GFP line[46]), *Nicotiana clevelandii*, *Spinacea oleracea*, *Chenopodium quinoa and Brassica napus* plants used for agroinfiltration or virus inoculation are grown on soil with 16 h light / 8 h darkness cycles in a greenhouse (at 21/18 °C). Infected grapevines originated from vineyards or from the INRAE-collection grown in individual pots to obtain two shoots of 180 cm, under natural light conditions.

Viral inoculations. Details about virus isolates/strains, inoculations, hosts and harvesting are listed in Supplementary Data 1. Briefly, PVX and TRV were inoculated from agrobacterium culture of infectious clones. GFLV (GT isolate), TMV, PCV, BBWV-1, CMV and TYMV were mechanically inoculated from purified virions, viral RNAs or infectious plasmid. TCV, TuMV, SoMV, AlMV, TBSV, BNYVV, ArMV (*C. quinoa*), TRSV, SLRSV (*C. quinoa*), GFLV (B844 isolate on *C. quinoa*), CNDV, CLRV, CPMV and RpRSV were inoculated by sap from infected tissues. TuYV was inoculated on *B. napus* by aphid transmission[47]. Grapevine rootstocks Kober 5BB were infected by GFLV (isolate B844) and ArMV by heterologous grafting and infected-vines were conserved in greenhouse. GRSPaV, GRVFV, GPGV, GRGV, GVB, TBRV, SLRSV, GLRaV-1, GLRaV-2 originate from vines in infected vineyards.

Virion purification. Viral particles of GFLV K30 and B844 isolates were purified from *C. quinoa* by sucrose gradient as described in[48,49].

Oligonucleotides. Oligonucleotides used in this study are listed in Supplementary Data 2.

RNA extraction from infected plants or purified viruses. Total RNA from infected leaves of *C. quinoa*, *N. benthamiana* and grapevine were extracted using the RNeasy plant Mini Kit (Qiagen) following the manufacturer's instructions except that the RLC lysis buffer was complemented with 25 mM DTT for herbaceous tissues or with 25 mM DTT and 1% (w/v) PVP40 (final concentrations) for grapevine tissues. Total RNA of Arabidopsis infected leaves was extracted using Tri-Reagent (Molecular Research Center). GFLV K30 and GFLV B844 genomic RNA was extracted from purified virions using phenol-chloroform. RNA concentrations were measured by spectrophotometry (Thermo Fisher scientific, Nanodrop 2000). RNA quality was checked by loading 200-400 ng total RNA on a 1% agarose gel.

3' RACE-seq library preparation. 1-5 µg of total RNA were denatured and ligated to 10 pmoles of a 5'-riboadenylated DNA oligonucleotide (3'-adapter RACE-seq, Supplementary Data 2a). For polyadenylated viruses, total RNA was denatured 5 min at 65°C. For non-polyadenylated viruses, total RNA was denatured 10 s at 95°C to increase ligation reaction efficiency. For all viruses, the ligation

reaction mixture contained 10 U of T4 ssRNA Ligase 1 (NEB) and 1X T4 RNA Ligase Reaction Buffer (NEB, 50 mM Tris-HCl pH 7.5, 10 mM MgCl2, 1 mM DTT) in a final volume of 50 µL. Samples with non-polyadenylated viruses were complemented with 6.25% (v/v).PEG 8000 (NEB). After 1h incubation at 37°C, surplus adapter and reagents were removed using Nucleospin RNA Clean-up columns (Macherey Nagel) following the manufacturer's instructions. cDNA synthesis was performed in 20 µl SuperScript™ IV RT buffer (Invitrogen) and contained 2 µg of purified ligated RNA, 50 pmol 3′-RT primer (Supplementary Data 2), 10 nmol dNTP, 0.1 µmol DTT, 40 U RNaseOUT (Invitrogen) and 200 U SuperScript™ IV reverse transcriptase. Reactions were incubated for 10 min at 50°C followed by an inactivation step at 80 °C for 10 min. Two nested PCR amplification rounds were performed using GoTaq® DNA Polymerase (Promega) using 1-2 µL cDNA or PCR1 as detailed in (Scheer et al. 2021). Gene-specific primers and TruSeq DNA PCR index are indicated in Supplementary Data 2a. PCR2 products were purified using one volume of magnetic beads (Sera-Mag™ Carboxylate-Modified Magnetic Beads, Cytiva). Libraries were paired-end sequenced on a MiSeq (v3 chemistry) with 41x111 bp cycle settings. To compensate for low diversity, Phix control v3 library was sequenced in parallel (30% of the flow cell).

5' RACE-seq library for GFLV. To analyze the 5' end sequence of the GFLV minus or plus strand, we set up a 5 'RACE-seq strategy which combines cDNA synthesis using the 5' RACE System for Rapid Amplification of cDNA End (Invitrogen™, v2.0) and Illumina sequencing. Briefly, 5 µg of total RNA extracted from Arabidopsis plants infected by GFLV (GT) were used to synthesize a GFLV specific cDNA following the manufacturer's instructions. Three rounds of PCR were then performed to amplify the 5' region using using GoTaq® DNA Polymerase (Promega) using 1-2 µL cDNA, PCR1 or PCR2. PCR cycles were as follow: a step at 94°C for 1 min; 25, 20 or 5 cycles (for PCR1, PCR2 or PCR3, respectively) at 94°C for 30 s, 55-65°C for 30 s and 72°C for 40 s; a final step at 72°C for 40 s. All used primers are listed in Supplementary Data 2a. All PCR3 products were purified using one volume of magnetic beads (Sera-Mag™ Carboxylate-Modified Magnetic Beads, Cytiva). Library was paired-end sequenced with MiSeq (v3 chemistry) with 41x111 bp cycle settings.

3' RACE-seq data processing. After initial data processing by the MiSeq Control Software v 2.6 (Illumina), Fastq files were analyzed by a homemade pipeline adapted from Scheer et al. 2021 composed of scripts using python (v2.7), biopython (v1.63)[50] and regex (v2.4) libraries. Reads with low quality bases (= < Q10) within the 15-base random sequence of the read 2 or within the 30 bases downstream of the delimiter sequence were filtered out. Sequences with identical nucleotides in 15-base random sequence were deduplicated. Next, 20 nucleotides sequences corresponding to nucleotides of the transcript that maps downstream the forward PCR2 primer (Supplementary Data 2a) were searched into reads 1 to identify the corresponding target viral RNA. One mismatch was tolerated. Matched reads 1 and their corresponding reads 2 were extracted for further analysis. Reads 2 that contain the delimiter sequence were selected and subsequently trimmed from their random and delimiter sequences. The rest of the analysis varied for known non-polyadenylated or polyadenylated viral RNA. For non-polyadenylated viral RNA the 30 nucleotide sequences downstream of the read 2 delimiter sequence were mapped to the corresponding reference sequence (Supplementary Data 1c), which goes from the first nucleotide of the transcript that maps the forward PCR2 primer to the end of the viral RNA

(see Supplementary Data 2a). Up to four mismatches were tolerated, with the exception of the first four nucleotides downstream of the mapping site that had to perfectly map. To map the 3' end position of reads 2 with untemplated tails, the sequences of the unmatched reads 2 were successively trimmed from their 3' end, with a one nucleotide trimming step, until they could be mapped to the reference sequence or until a maximum of 30 nucleotide has been removed. For each successfully mapped read 2, untemplated nucleotides at the 3' end were extracted and annotated according to their nucleotide tail composition as: U-tails (*i.e.* composed of only Us), U-rich tails (*i.e.* composed of a majority of Us, at least 70%), A- tails (*i.e.* composed of. only As), A-rich tails (*i.e.* composed of a majority of As, at least 70%), C- tails (*i.e.* composed of only Cs), C-rich tails (*i.e.* composed of a majority of Cs, at least 70%), G- tails (*i.e.* composed of only Gs) and G-rich tails (i.e. composed of a majority of Gs, at least 70%). For polyadenylated viral RNAs, the analysis was divided into two steps. The aim of the first step was to identify the position of viral RNA 3' extremities and to detect untemplated nucleotides and was performed as described above for non-polyadenylated viral RNA. Of note, this first step was skipped for representative viruses of the *Secoviridae* family, as the complete reference sequence was missing for most of them (results shown in Fig. 8 and Fig. S6). The aim of the second step was to analyze long RNA poly(A) tails. Sequencing of long homopolymeric stretches causes a rapid decrease of sequencing quality, making it impossible to exactly map the 3' end of RNA with long poly(A). We thus looked for long T stretches of at least 8 Ts (*i. e.* the complementary sequence of a poly(A)) in read 2 that failed to map the reference sequence. Poly(A) tails were searched with the constraint that it must begin in the first 30 cycles, which means that the maximal length of the added 3' end modification is limited to 29 nucleotides. Finally, results from step 1 and 2 were compiled and the 3' tail were analyzed and annotated based on their composition as: non modified poly(A) tails, uridylated tails (poly(A) + U or poly(U)), cytidinylated tails (poly(A) + C or poly(C)) and guanidylated tails (poly(A) + G or poly(G)). For GLRaV-2 RNA, A-rich tails ending by U were annotated as uridylated. Python and bash source code are available in github (github link).

Distribution profiles shown in Fig. 6 and in Supplementary Data 4 display the percentages of sequences according to poly(A) tail sizes calculated for non-modified (light gray) or uridylated (turquoise) poly(A) tails from 1 to 89 nucleotides. Percentages were calculated using the total number of reads as denominator. Heatmaps shown in Fig. 1, 3 and 4 and in Supplementary Fig. 2 and 3 display the frequency at each 3' end position for all reads (purple color scale), non-tailed reads (blue color scale), U-tailed reads (orange color scale) or U-rich tailed reads (green color scales). Frequencies were calculated using the total number of reads as denominator.

5' RACE-seq data processing. Most steps of 5' RACE-seq data processing are common to 3' RACE-seq data analyses, except for the few steps detailed below. First, reads were deduplicated using a random sequence that contains 8 bases instead of 15. Second, the initial quality filter initial has been raised compared to 3' RACE-seq and reads with low quality bases (= < Q20) within the 8-base random sequence of the read 2 or within the 30 bases downstream of the delimiter sequence were filtered out. Indeed, the read 2 sequence contains an additional G homopolymeric stretches (a C-tail is added at the 5' end of the cDNA) causing a decrease of sequencing quality. This G homopolymeric stretches was removed after the deduplication step together with the random, delimiter sequence. Finally, we searched

for long poly(U) tails, instead of poly(A), and 5' tails were annotated as non-modified poly(U) tails, adenylated tails (poly(U) + A or poly(A)) and cytidinylated tails (poly(U) + C or poly(C)). One note, guanidylated tails (poly(U) + G or poly(G)) could not be detected as the read2 sequence contained G homopolymeric stretches due to library preparation procedure. Python and bash source code are available in github (https://github.com/hzuber67/RACEseq_virus).

Statistics and reproducibility. All plots were generated using R (v. 3.6.1) and the R package ggplot2 (v. 3.3.5) on RStudio (v. 1.4.1106). To compare uridylation percentages of viral RNAs (Fig. 4a and 6a) and the frequency of each nucleotide before tail (Fig. 4e) across genotypes, we used the R package car (v3.0-5) applying a generalized linear model for proportions with a quasibinomial distribution. The multcomp package (1.4-19) with Tukey contrasts was used for multiple comparison post hoc tests and the calculation of adjusted p-values. Detailed results of the statistical analysis are provided in Supplementary Data 4. For all statistical analyses, a p-value of 0.05 is defined as threshold of significance. The number of independent biological replicates is indicated in each figure legend.

Biological material availability. All biological materials used in this study are available from the authors.

Data availability. NGS datasets generated during this study have been deposited in NCBI's Gene Expression Omnibus[51] and are accessible through GEO Series accession number GSEXXX (link). GEO Series accession numbers for individual datasets are GSEXXX (link). Web links for associated raw data are indicated in each figure legends. Source data for all figures are included as Supplemental Data:

https://data.mendeley.com/datasets/k3njd5yfj3/draft?a=bd2e49d5-ac42-4b84-befb-23e6df7e2011

Code availability. Bioinformatic pipelines including python and bash source code for 3'RACE-seq are available in Github (https://github.com/hzuber67/RACEseq_virus)

References

1. Calil, I. P. & Fontes, E. P. B. Plant immunity against viruses: antiviral immune receptors in focus. Ann Bot 119, 711–723 (2017).
2. Le Pen, J. et al. Terminal uridylyltransferases target RNA viruses as part of the innate immune system. Nat. Struct. Mol. Biol. 25, 778–786 (2018).
3. Yu, S. & Kim, V. N. A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. Nat Rev Mol Cell Biol 21, 542–556 (2020).
4. Zigáčková, D. & Vaňáčová, Š. The role of 3' end uridylation in RNA metabolism and cellular physiology. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 373, (2018).
5. Scheer, H., Zuber, H., De Almeida, C. & Gagliardi, D. Uridylation earmarks mRNAs for degradation… and more. Trends in Genetics 32, 607–619 (2016).
6. De Almeida, C., Scheer, H., Zuber, H. & Gagliardi, D. RNA uridylation: a key post-transcriptional modification shaping the coding and non-coding transcriptome. WIREs RNA e1440 (2018) doi:10.1002/wrna.1440.
7. Warkocki, Z., Liudkovska, V., Gewartowska, O., Mroczek, S. & Dziembowski, A. Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 373, (2018).
8. Huo, Y. et al. Widespread 3′-end uridylation in eukaryotic RNA viruses. Sci Rep 6, (2016).
9. Menzel, W. & Vetten, H. J. Complete nucleotide sequence of an isolate of the Anthriscus strain of Parsnip yellow fleck virus. Arch Virol 153, 2173–2175 (2008).
10. Hulo, C. et al. ViralZone: a knowledge resource to understand virus diversity. Nucleic Acids Res 39, D576-582 (2011).
11. Zhu, H. Y., Ling, K. S., Goszczynski, D. E., McFerson, J. R. & Gonsalves, D. Nucleotide sequence and genome organization of grapevine leafroll-associated virus-2 are similar to beet yellows virus, the closterovirus type member. J Gen Virol 79 (Pt 5), 1289–1298 (1998).

12. Liu, Y.-P., Peremyslov, V. V., Medina, V. & Dolja, V. V. Tandem leader proteases of Grapevine leafroll-associated virus-2: Host-specific functions in the infection cycle. Virology 383, 291–299 (2009).

13. Dreher, T. W. Viral tRNAs and tRNA-like structures. Wiley Interdiscip Rev RNA 1, 402–414 (2010).

14. Olsthoorn, R. C., Mertens, S., Brederode, F. T. & Bol, J. F. A conformational switch at the 3' end of a plant virus RNA regulates viral replication. EMBO J 18, 4856–4864 (1999).

15. Yuan, X., Shi, K. & Simon, A. E. A local, interactive network of 3' RNA elements supports translation and replication of Turnip crinkle virus. J Virol 86, 4065–4081 (2012).

16. Simon, A. E. 3'UTRs of carmoviruses. Virus Res 206, 27–36 (2015).

17. McCormack, J. C. et al. Structural domains within the 3' untranslated region of Turnip crinkle virus. J Virol 82, 8706–8720 (2008).

18. Le, M.-T. et al. Folding behavior of a T-shaped, ribosome-binding translation enhancer implicated in a wide-spread conformational switch. eLife 6, e22883 (2017).

19. Sement, F. M. et al. Uridylation prevents 3' trimming of oligoadenylated mRNAs. Nucleic Acids Res. 41, 7115–7127 (2013).

20. Ren, G., Chen, X. & Yu, B. Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Curr. Biol. 22, 695–700 (2012).

21. Zhao, Y. et al. The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. Curr. Biol. 22, 689–694 (2012).

22. Tu, B. et al. Distinct and Cooperative Activities of HESO1 and URT1 Nucleotidyl Transferases in MicroRNA Turnover in Arabidopsis. PLoS Genet 11, e1005119 (2015).

23. Zuber, H., Scheer, H., Joly, A.-C. & Gagliardi, D. Respective Contributions of URT1 and HESO1 to the Uridylation of 5′ Fragments Produced From RISC-Cleaved mRNAs. Front. Plant Sci. 9, (2018).

24. Zuber, H. et al. Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis. Cell Rep 14, 2707–2717 (2016).

25. Scheer, H. et al. The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis. Nat Commun 12, 1298 (2021).

26. Zhang, W., Murphy, C. & Sieburth, L. E. Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. Proc. Natl. Acad. Sci. U.S.A. 107, 15981–15985 (2010).

27. Malecki, M. et al. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J. 32, 1842–1854 (2013).

28. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. Nature 514, 252–256 (2014).

29. Ritzenthaler, C. et al. Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. Journal of General Virology 72, 2357–2365.

30. Serghini, M. A. et al. RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location. J Gen Virol 71 (Pt 7), 1433–1441 (1990).

31. Arnold, J. J. & Cameron, C. E. Poliovirus RNA-dependent RNA Polymerase (3Dpol) Is Sufficient for Template Switchingin Vitro *. Journal of Biological Chemistry 274, 2706–2716 (1999).

32. Neufeld, K. L., Galarza, J. M., Richards, O. C., Summers, D. F. & Ehrenfeld, E. Identification of terminal adenylyl transferase activity of the poliovirus polymerase 3Dpol. J Virol 68, 5811–5818 (1994).

33. Ranjith-Kumar, C. T. et al. Terminal nucleotidyl transferase activity of recombinant Flaviviridae RNA-dependent RNA polymerases: implication for viral RNA synthesis. J. Virol. 75, 8615–8623 (2001).

34. Poranen, M. M., Koivunen, M. R. L. & Bamford, D. H. Nontemplated terminal nucleotidyltransferase activity of double-stranded RNA bacteriophage phi6 RNA-dependent RNA polymerase. J Virol 82, 9254–9264 (2008).

35. Fullerton, S. W. B. et al. Structural and functional characterization of sapovirus RNA-dependent RNA polymerase. J Virol 81, 1858–1871 (2007).

36. Rohayem, J. et al. Characterization of norovirus 3Dpol RNA-dependent RNA polymerase activity and initiation of RNA synthesis. J Gen Virol 87, 2621–2630 (2006).

37. Tomar, S., Hardy, R. W., Smith, J. L. & Kuhn, R. J. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. J Virol 80, 9962–9969 (2006).

38. Wu, W. et al. Flock house virus RNA polymerase initiates RNA synthesis de novo and possesses a terminal nucleotidyl transferase activity. PLoS One 9, e86876 (2014).

39. Wang, Z. et al. Characterization of a nodavirus replicase revealed a de novo initiation mechanism of RNA synthesis and terminal nucleotidyltransferase activity. J Biol Chem 288, 30785–30801 (2013).

40. Behrens, S. E., Tomei, L. & De Francesco, R. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. EMBO J 15, 12–22 (1996).

41. Smallwood, S. & Moyer, S. A. Promoter analysis of the vesicular stomatitis virus RNA polymerase. Virology 192, 254–263 (1993).

42. Tvarogová, J. et al. Identification and Characterization of a Human Coronavirus 229E Nonstructural Protein 8-Associated RNA 3'-Terminal Adenylyltransferase Activity. J Virol 93, e00291-19 (2019).

43. Fuchs, M. & Lemaire, O. Novel Approaches for Viral Disease Management. in Grapevine Viruses: Molecular Biology, Diagnostics and Management (eds. Meng, B., Martelli, G. P., Golino, D. A. & Fuchs, M.) 599–621 (Springer International Publishing, 2017). doi:10.1007/978-3-319-57706-7_29.

44. Mannini, F. & Digiaro, M. The Effects of Viruses and Viral Diseases on Grapes and Wine. in Grapevine Viruses: Molecular Biology, Diagnostics and Management (eds. Meng, B., Martelli, G. P., Golino, D. A. & Fuchs, M.) 453–482 (Springer International Publishing, 2017). doi:10.1007/978-3-319-57706-7_23.

45. Schmitt-Keichinger, C., Hemmer, C., Berthold, F. & Ritzenthaler, C. Molecular, Cellular, and Structural Biology of Grapevine fanleaf virus. in Grapevine Viruses: Molecular Biology, Diagnostics and Management (eds. Meng, B., Martelli, G. P., Golino, D. A. & Fuchs, M.) 83–107 (Springer International Publishing, 2017). doi:10.1007/978-3-319-57706-7_4.

46. Monsion, B. et al. Efficient Detection of Long dsRNA in Vitro and in Vivo Using the dsRNA Binding Domain from FHV B2 Protein. Frontiers in plant science 9, (2018).

47. Leiser, R. M. et al. Agroinfection as an alternative to insects for infecting plants with beet western yellows luteovirus. Proceedings of the National Academy of Sciences of the United States of America 89, 9136–9140 (1992).

48. Pinck, L., Fuchs, M., Pinck, M., Ravelonandro, M. & Walter, B. Y. 1988. A Satellite RNA in Grapevine Fanleaf Virus Strain F13. Journal of General Virology 69, 233–239.

49. Schellenberger, P. et al. Structural insights into viral determinants of nematode mediated Grapevine fanleaf virus transmission. PLoS Pathog. 7, e1002034 (2011).

50. Cock, P. J. A. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423 (2009).

51. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30, 207–210 (2002).

Author Contributions

ACJ and SG performed most of the experiments; HZ designed all bioinformatic analyses that were conducted by HZ and ACJ; HZ and ACJ produced all figures and associated datasets, except for the phylogenetic analysis of *Secoviridae* viruses performed by JMH; ACJ, SG, JMH, DaGa, EV, OL, HZ and DoGa discussed and analyzed data; SK performed the MiSeq runs; HZ and DoGa supervised the project; DoGa initiated the project and acquired funding.

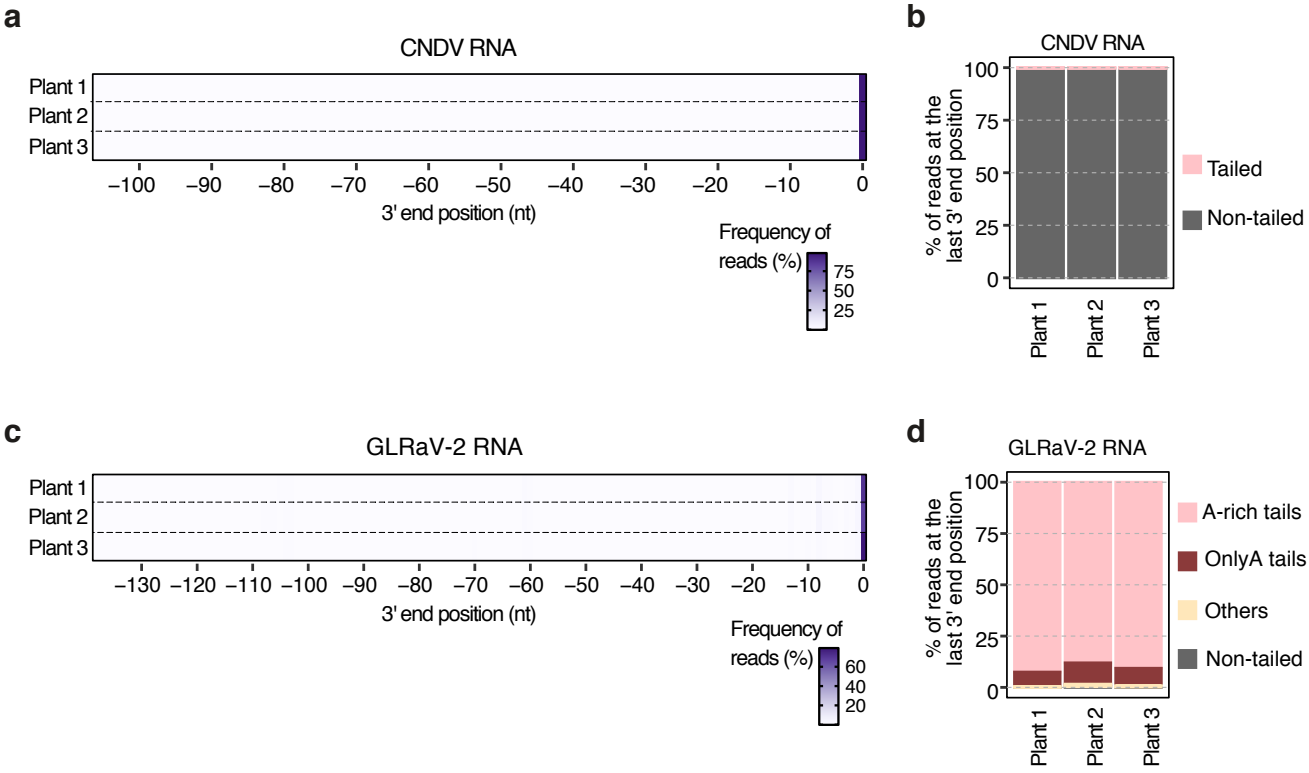
Acknowledgements

The authors gratefully acknowledge Salah Bouzoubaah, Véronique Brault, Marc Fuchs, David Gilmer, Anthony Gobert, Manfred Heinlein, Jean-Sébastien Reynard, Christophe Ritzenthaler, Corinne Schmitt-Keichinger, Claire Villeroy and Véronique Ziegler-Graff for providing viruses or infected plant material analyzed in this study. The authors also thank Anthony Gobert for sharing the experimental procedure for ligation of structured RNA and expressed their gratitude to David Gilmer and Corinne Schmitt-Keichinger for helpful discussions. This research was funded by the Centre National de la Recherche Scientifique (CNRS) and by a research grant from the French National Research Agency ANR-20-CE20-0010 coordinated by D.G. Some of the sequencing experiments were supported by a funding from the state managed by the French National Research Agency as part of the “Investments for the Future” program under the framework of the LABEX: ANR-10-LABX-0036_NETRNA and ANR-17-EURE-0023.

Declaration of Interests

The authors declare no competing interests.

c. Data related to the article



e Examples of A-rich tails observed for GLRaV-2 RNA

AAAAAAAAAUAAAAAAAAAAAAAAAGAAGC
AUAAAAAAAAAAAAAAAAAAAAAAAAAAAG
AAAAAAAAUAAAUAAAAAAAAAAAAAAAGA
AAAAAAAAUAUAAAAAAAAAAAAAAAGAU
AAAAAAAAUAUAAAAAAAAAAAAAAAAA
AAAAAAAAUAAAAAAAAAAAAGAAGCC
AAAAAAAAUAAAAAAAAAUUUUUAAAAU
AAAAAAAAAUAUAAAAAAAGAAAA
AUAUUUAAAAAAAAAAAAAUGAAG
AAAAAAAUAAAAAUUUAAAAAUUU
AAAAAAAUAAAAAAAUUUUCU
AAAAAAAUAUAAAAAAAAG
AAAAAAAAAAAAUAAAAGU
AAAAAAAAAAAUUUUU
AAAAAAAGAGGC
AAAAAUUUAA

f Sequence logo for GLRaV-2 RNA tails that contain G

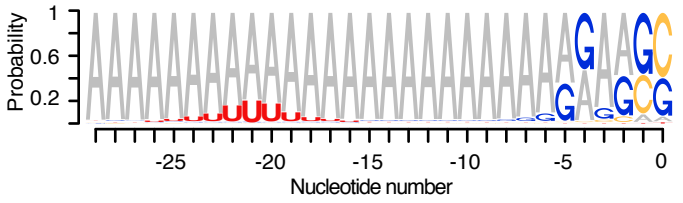


Figure 1 3' terminal features of CNDV and GLRaV-2 RNAs. a, c High-resolution mapping of 3' ends for CNDV (a) and GLRaV-2 (c) RNA. Heatmaps show frequencies of reads mapped to the indicated position for three infected plants. Position 0 corresponds to the 3' end of the full-length viral RNA. **b, d** Percentage of tailed vs non-tailed reads found at the last 3' end position for CNDV (b) and GLRaV-2 (d) RNA. **e** Examples of A-rich tails added to GLRaV-2 RNA. **f** Sequence logo generated from GLRaV-2 RNA tails containing G. 42 to 66% of tail sequences contain G. Plant hosts are indicated in Supplementary Data 1a.

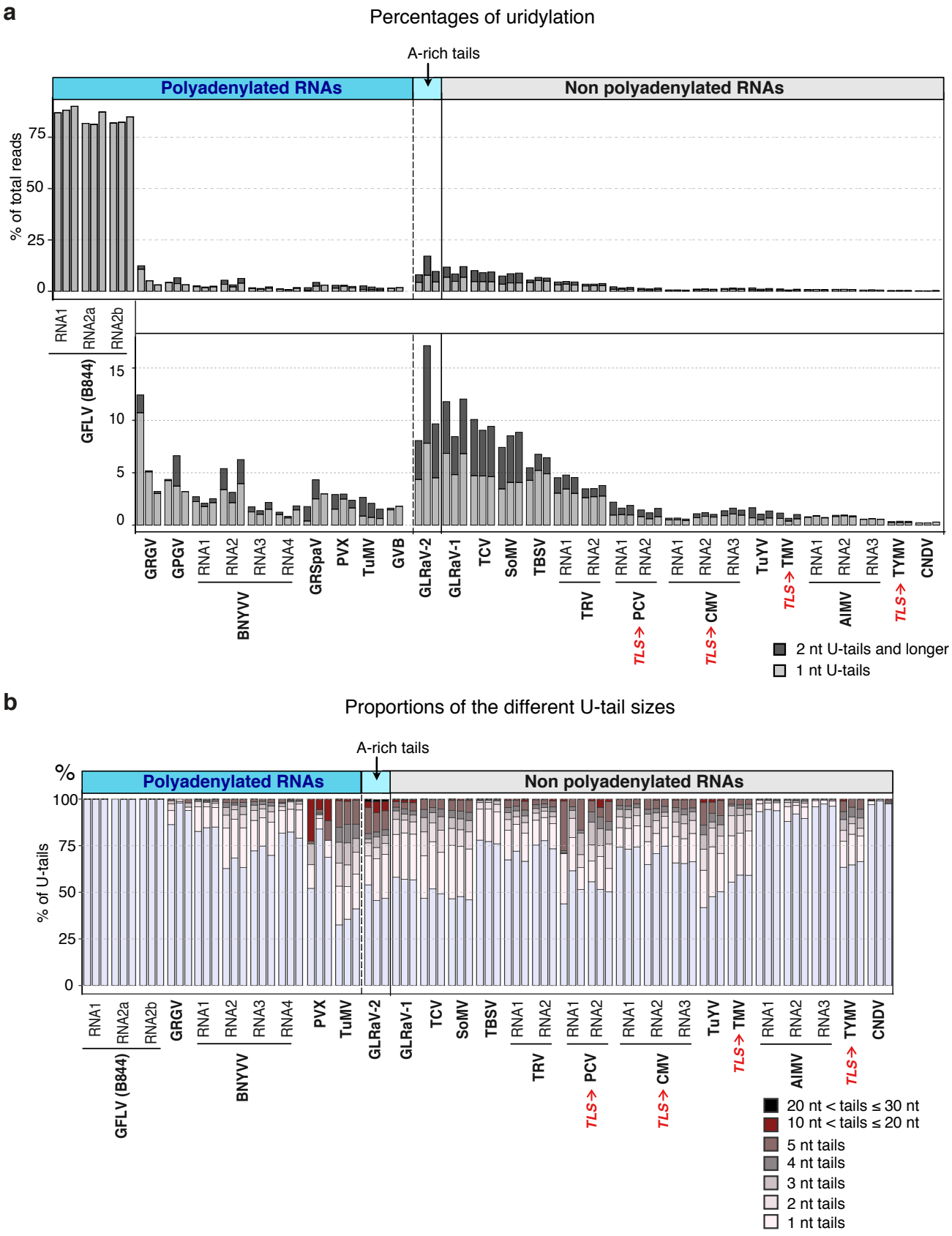


Figure 2 U-tailing of ss(+) RNA phytoviruses is widespread and diverse. a Uridylation percentages of phytoviral RNAs. Each bar represents an infected plant (n=2 for GVB and n=3 for all other viruses). The percentages of long (> 1 U) and 1 U-tails are indicated by dark gray and light gray, respectively. **b** Proportion of the different U-tail sizes from 1 to 30 nt. Percentages were calculated using the number of U-tailed reads as denominator. U-tail sizes are indicated by a color gradient from light purple, for 1U, to black for 20 to 30 Us. Only viral RNAs with at least 50 U-tailed reads for each replicate are shown. Plant hosts are indicated in Supplementary Data 1a.

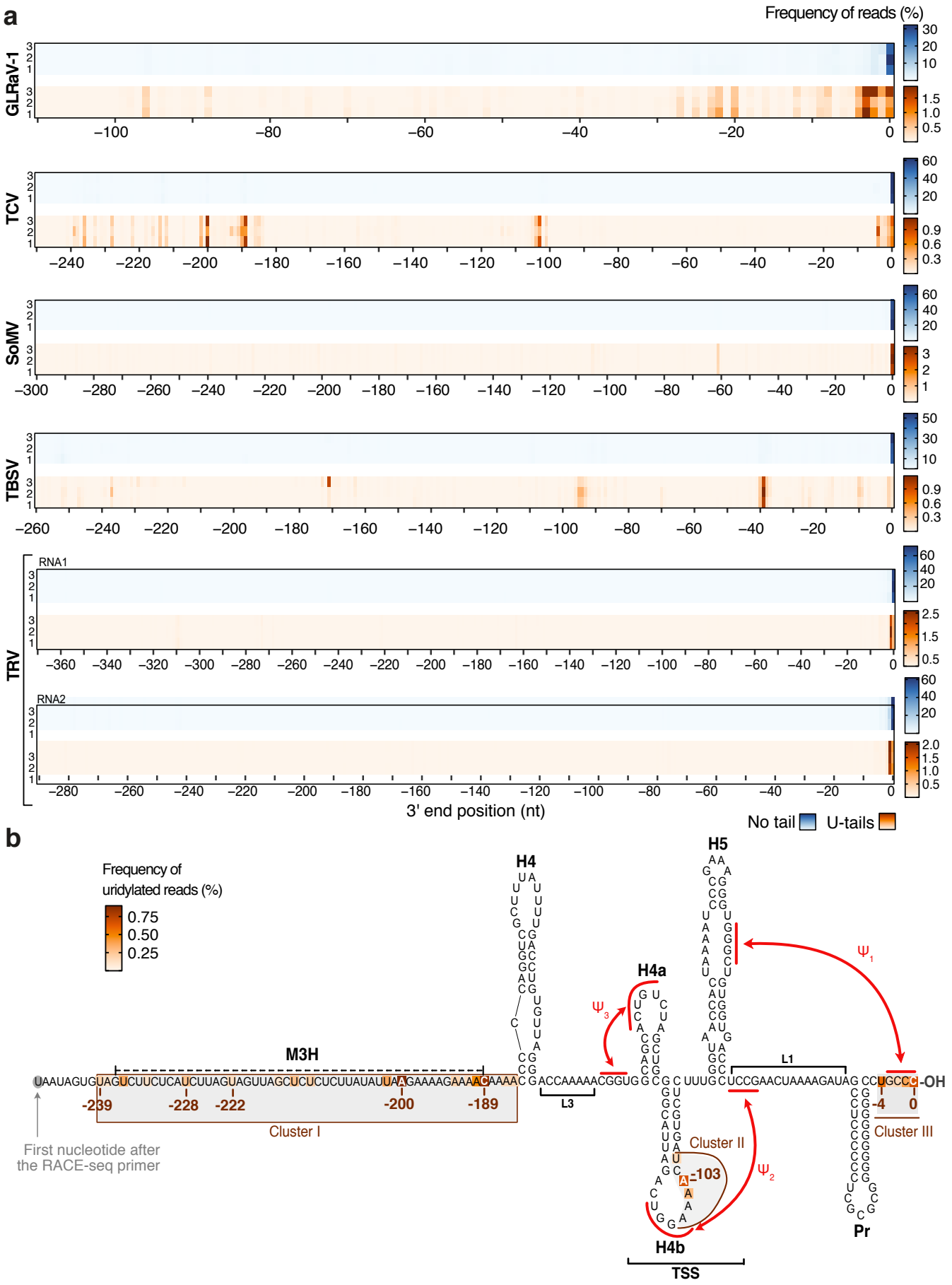


Figure 3: Uridylation of degradation intermediates reveals patterns of ribonucleolytic attacks. a High-resolution mapping of RNA 3' ends for a selection of non-polyadenylated viral RNAs (indicated on the left). Frequencies of reads at each 3' end position are shown by a blue color scale for non-tailed reads and an orange color scale for U-tailed reads. Frequencies were calculated using the total number of reads as denominator. Position 0 corresponds to the 3' end of the full-length viral RNA. For each virus, three infected plants were analyzed. **b** Secondary structure and tertiary interactions in the 3′ UTR of TCV RNA according to (McCormack et al., 2008; Simon and Miller, 2015). The TCV 3' UTR contains one weak (M3H) and five stable hairpins (H4, H4a, H4b, H5 and Pr) as well as three H-type pseudoknots (Y1, Y2, Y3) shown as red arrows. The frequency of uridylated 3' ends detected in infected *N. benthamiana* plants is indicated by colored rectangles for each uridylation site. The three detected clusters of uridylation sites are highlighted in yellow. Plant hosts are indicated in Supplementary Data 1a.

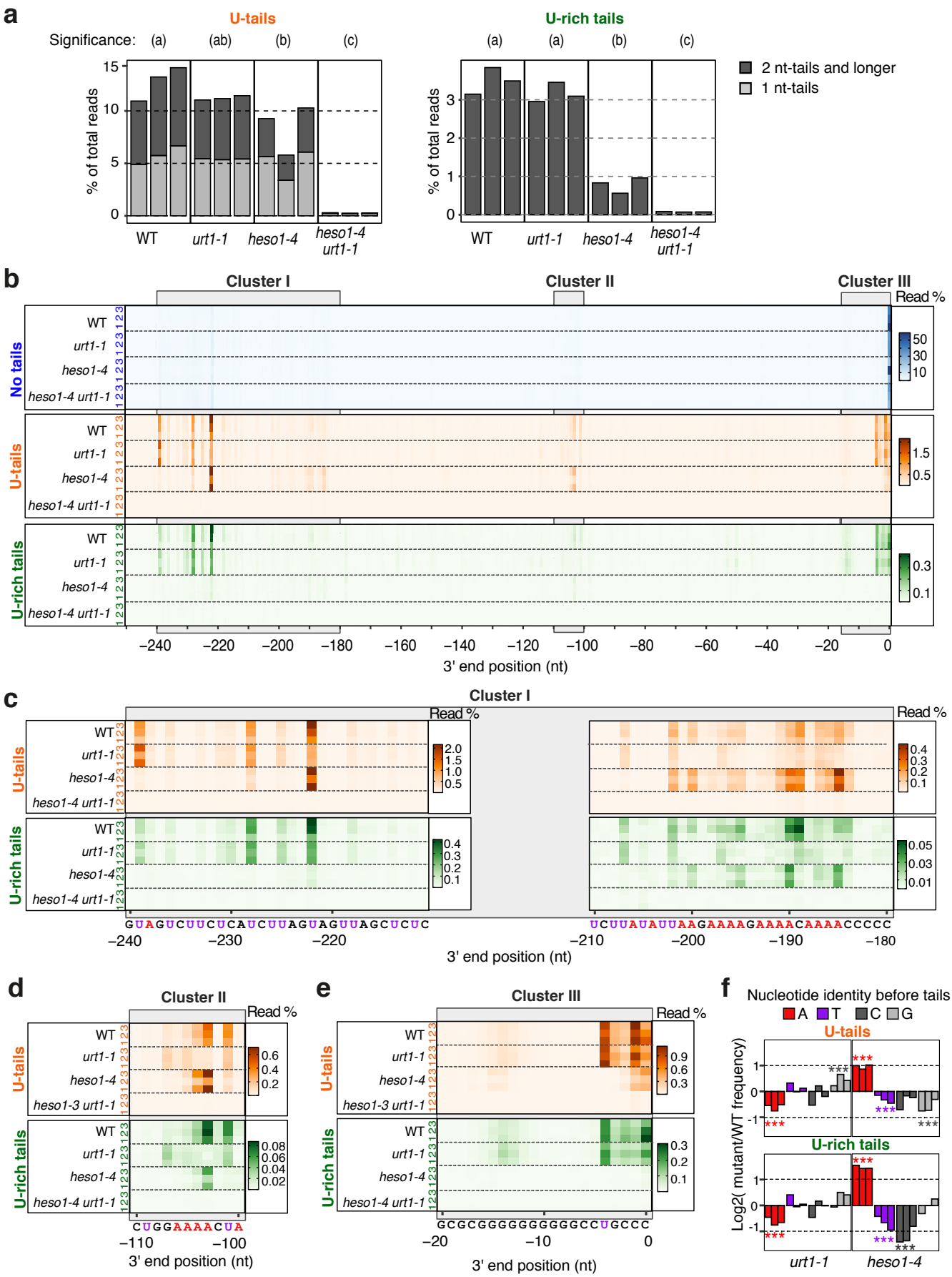


Figure 4: Contribution of the Arabidopsis TUTases in the uridylation of TCV RNAs. a Uridylation percentages of TCV RNA in infected WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* plants. Percentages are shown for tails containing only Us (U-tail, left panel) or a majority of Us (U-rich, right panel). Each bar represents an individual plant (n=3). The percentages of long (> 1 U) and 1 U-tails are indicated by dark gray and light gray, respectively. Significantly different values ($p < 0.05$) are labelled by different letters (generalized linear model for proportion, quasibinomial distribution) **b-e** High resolution mapping of TCV RNA 3' ends in infected WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* plants. Frequencies were calculated using the total number of reads as denominator. Frequencies of non-tailed, U-tailed and U-rich-tailed reads at each 3' end position are shown by blue, orange and green color scales, respectively. Position 0 corresponds to the 3' end of full-length TCV RNA. A close-up view is shown for the three detected clusters of uridylation in c-e. **f** Relative frequency, compared to WT, of the last nucleotide before U-tails and U-rich tails in *urt1-1* and *heso1-4* mutants. Stars represent significant statistical p-value (generalized linear model for proportion, quasibinomial distribution, n=3) with $p < 0.001$.

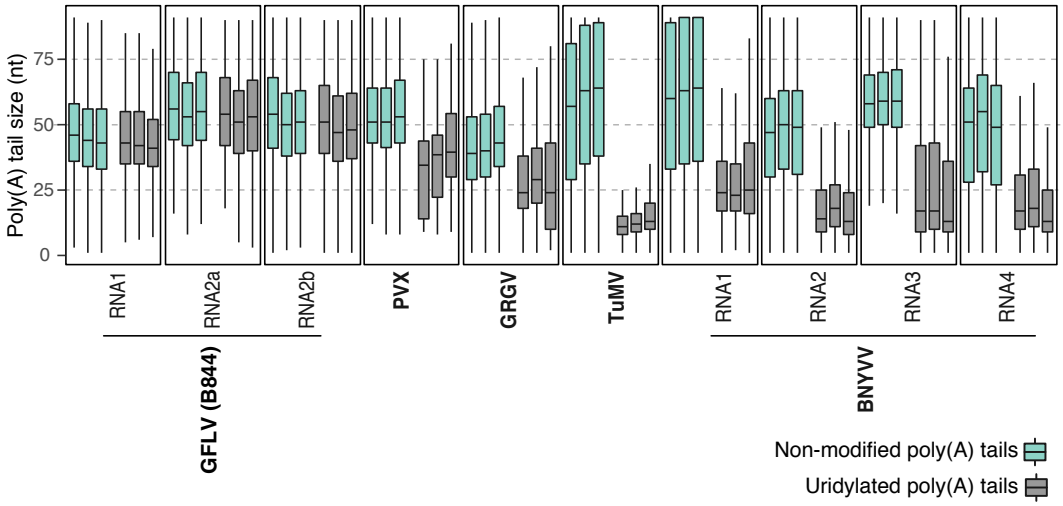


Figure 5 Boxplot analysis comparing the size of non-modified poly(A) tails vs uridylated poly(A) tails for polyadenylated viral RNAs. Each boxplot represents an infected Arabidopsis plant (n=3) and displays the median, first and third quartiles (lower and upper hinges), the largest value within 1.5 times the interquartile range above the upper hinge (upper whisker) and the smallest value within 1.5 times the interquartile range below the lower hinge (lower whiskers). Only viral RNAs with at least 50 U-tailed reads for each replicate are shown. Plant hosts are indicated in Supplementary Data 1a.

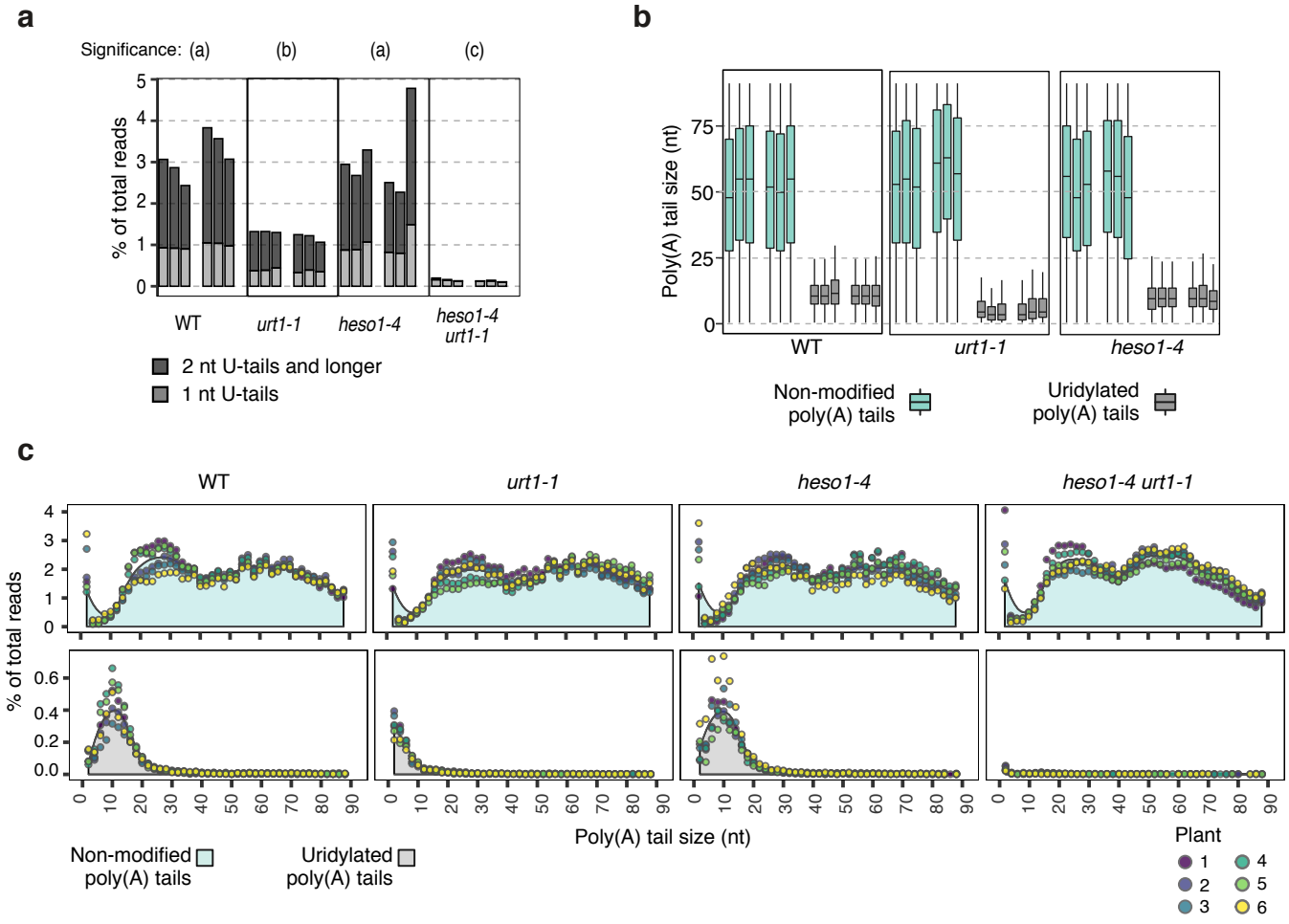


Figure 6 Both URT1 and HESO1 contribute to the uridylation of TuMV RNA. a Percentages of uridylation of TuMV RNA in infected WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* Arabidopsis plants. Each bar represents an infected plant (n=6). Percentages of long (> 1 U) and 1 U-tails are indicated by dark gray and light gray, respectively. Significantly different values ($p<0.05$) are labelled by different letters (generalized linear model for proportion, quasibinomial distribution). **b** Boxplot analysis comparing non-modified poly(A) tails (turquoise) vs uridylated poly(A) tails (gray). Each boxplot represents an infected plant (n=6) and displays the median, first and third quartiles (lower and upper hinges), the largest value within 1.5 times the interquartile range above the upper hinge (upper whisker) and the smallest value within 1.5 times the interquartile range below the lower hinge (lower whiskers). **c** Distribution of poly(A) tail sizes of non-tailed (turquoise) or uridylated (gray) viral RNAs for WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* plants infected by TuMV. Percentages were calculated using the total number of sequences with tails from 1 to 89 nucleotides as denominator. Individual points are color-coded for each of the six replicates. The gray area indicates the average of all replicates.

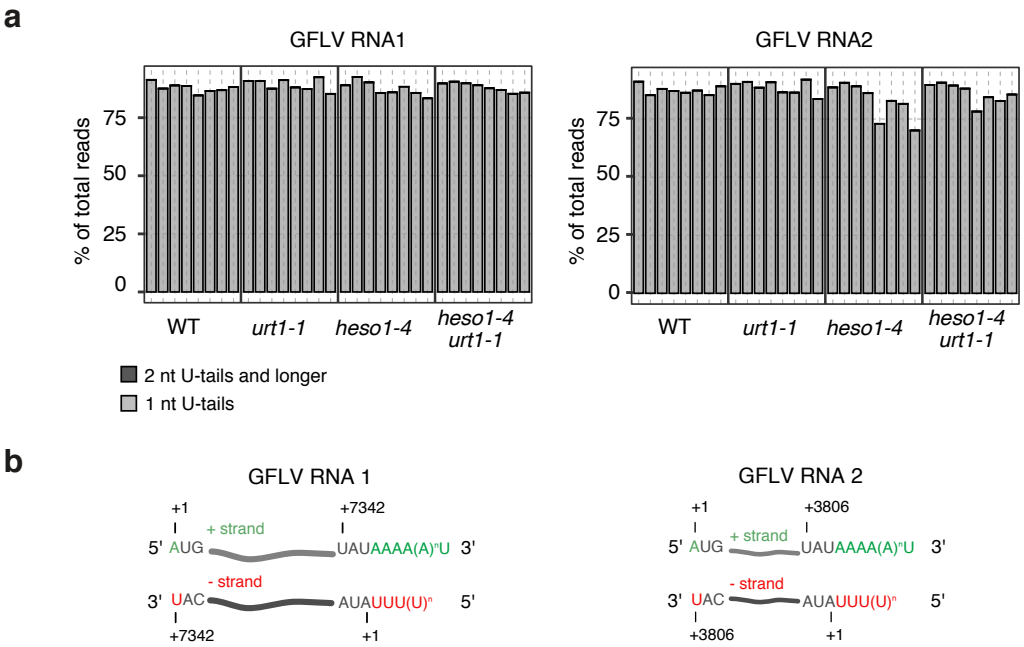


Figure 7 Arabidopsis TUTases are not required to maintain uridylation of GFLV RNAs. a Uridylation percentages of GFLV RNAs (isolate GT) in infected WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* Arabidopsis plants. Each bar represents an infected plant (n=8). The percentages of long (> 1 U) and 1 U-tails are indicated by dark gray and light gray, respectively. **c** Diagram illustrating the 3' and 5' extremities of the GFLV RNA plus and minus strands deduced from 3' and 5'RACE-seq results (shown in Supplementary Fig. 5).

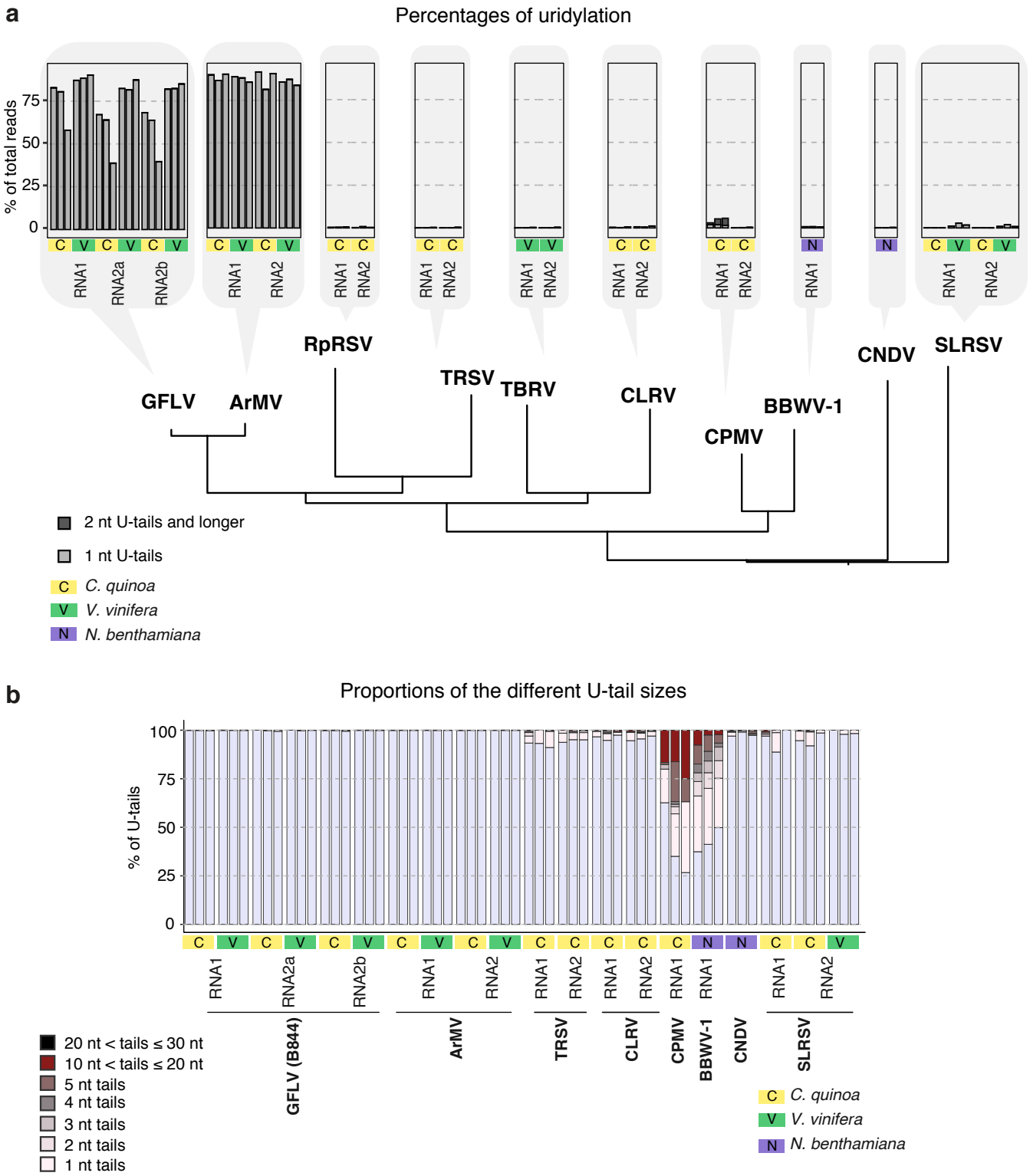


Figure 8 High uridylation levels are restricted to GFLV and ArMV. a Uridylation percentages among the *Secoviridae* family. Each bar represents an infected plant, the host is indicated below each bar. The percentages of long (> 1 U) and 1 U-tails are indicated by dark gray and light gray, respectively. The diagram below barplots illustrates the phylogenetic distances between the *Secoviridae* viruses analyzed in this study. The phylogenetic tree is shown in Supplementary Fig. 7b. **b** Proportion of the different U-tail sizes from 1 to 30 nt. The percentages were calculated using the number of U-tails as denominator. U tail sizes are indicated by a color gradient from light purple for 1-nt U-tails to black for 20 to 30-nt U-tails. Only viral RNAs with at least 50 U-tailed reads for each replicate are shown.

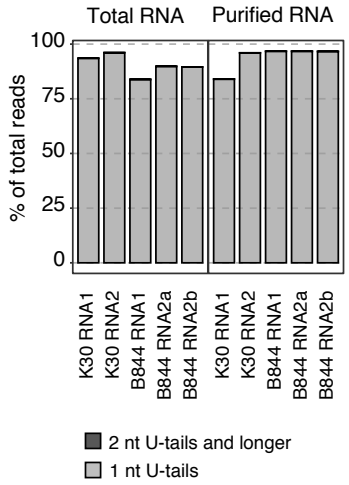


Figure 9 Uridylation is a genomic feature of encapsidated GFLV RNAs. Uridylation frequencies of GFLV RNAs (K30 and B844 isolates) for total RNA of *C. quinoa* infected plants or for purified RNA from GFLV virions. For each of the two GFLV isolates, one replicate was analyzed.

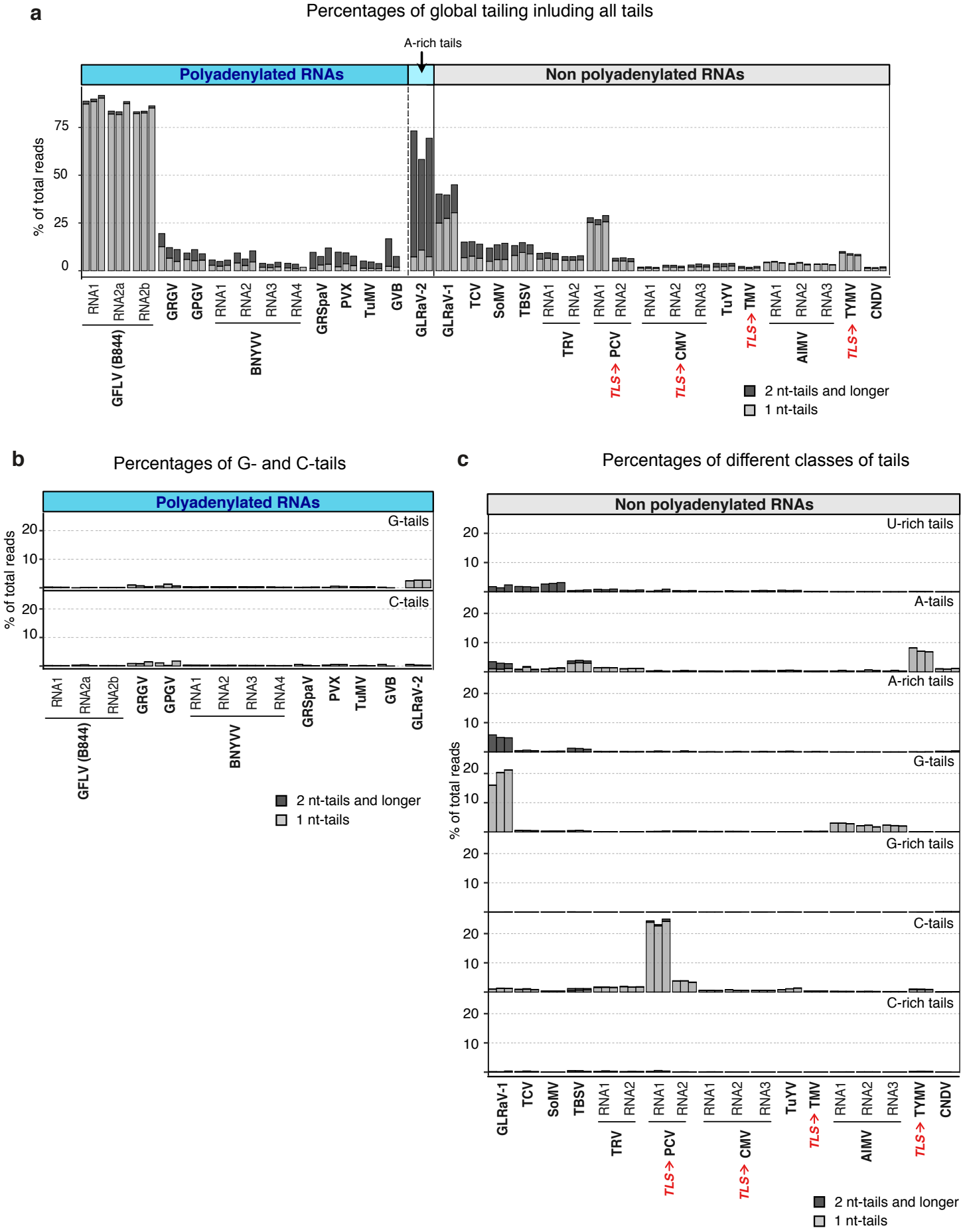


Figure S1: Diversity of phytoviral RNA tailing among ss(+) RNA phytoviruses. a. Global tailing percentages of phytoviral RNAs including all nucleotide additions, A, U, C, G and mixed tails. **b.** Percentages of G- and C-tails detected in polyadenylated viral RNAs. **c.** Percentages of different classes of tails added to the 3' ends of non-polyadenylated viral RNAs. The percentages of long (> 1 nt) and 1 nt-tails are indicated by dark gray and light gray, respectively. Each bar represents an infected plant (n=3). Plant hosts are indicated in Supplementary Data 1a.

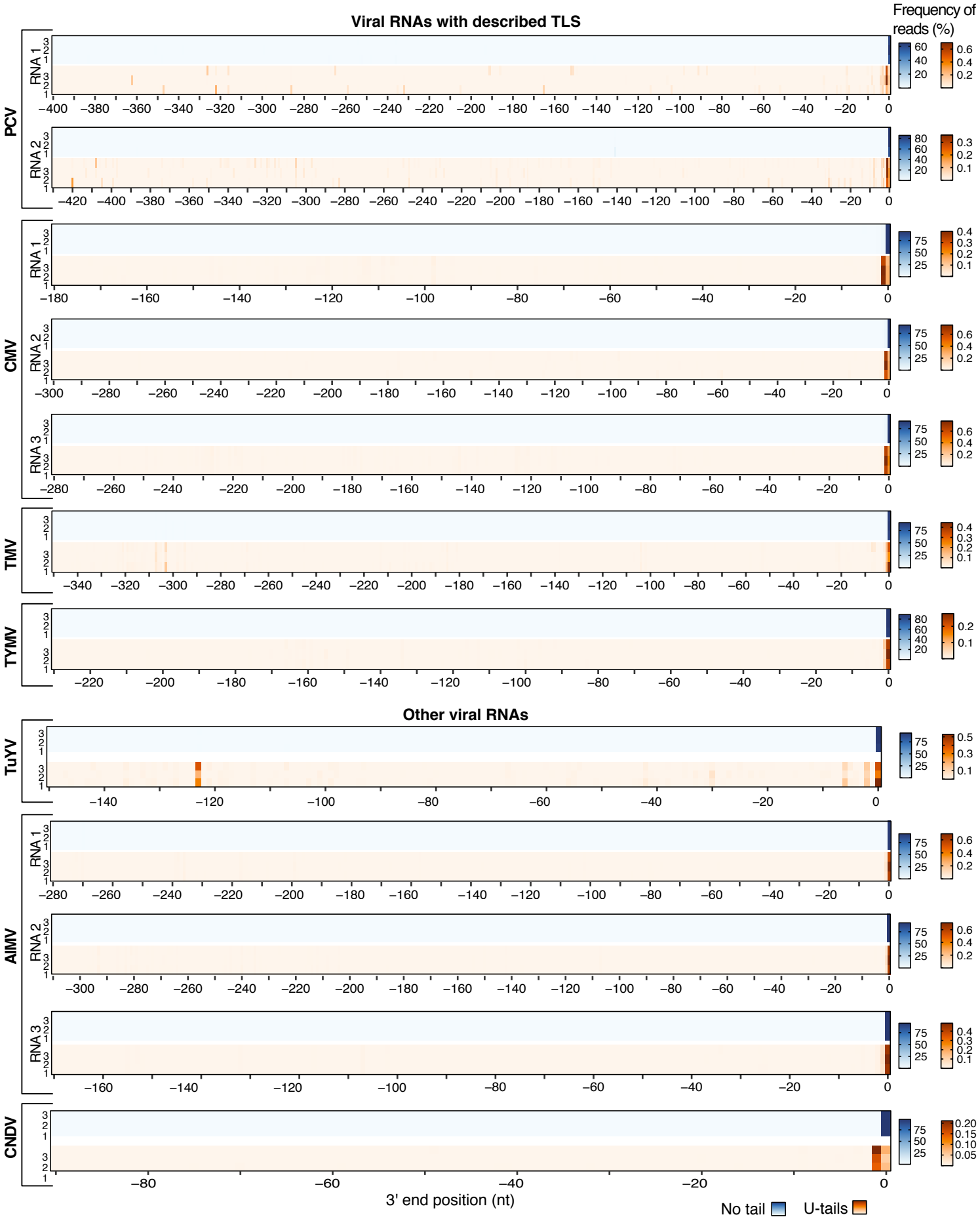


Figure S2 Phytoviral RNAs are mostly uridylated at their 3' ends. High resolution mapping of RNA 3' ends for non-polyadenylated virus with or without known 3' TLS structures. Frequencies of reads at each 3' end position are shown by a color scale for non-tailed reads (blue) and U-tailed reads (orange). Frequencies were calculated using the total number of reads as denominator. The position 0 corresponds to the known 3' end. For each virus, three infected plants were analyzed. Plant hosts are indicated in Supplementary Data 1a.

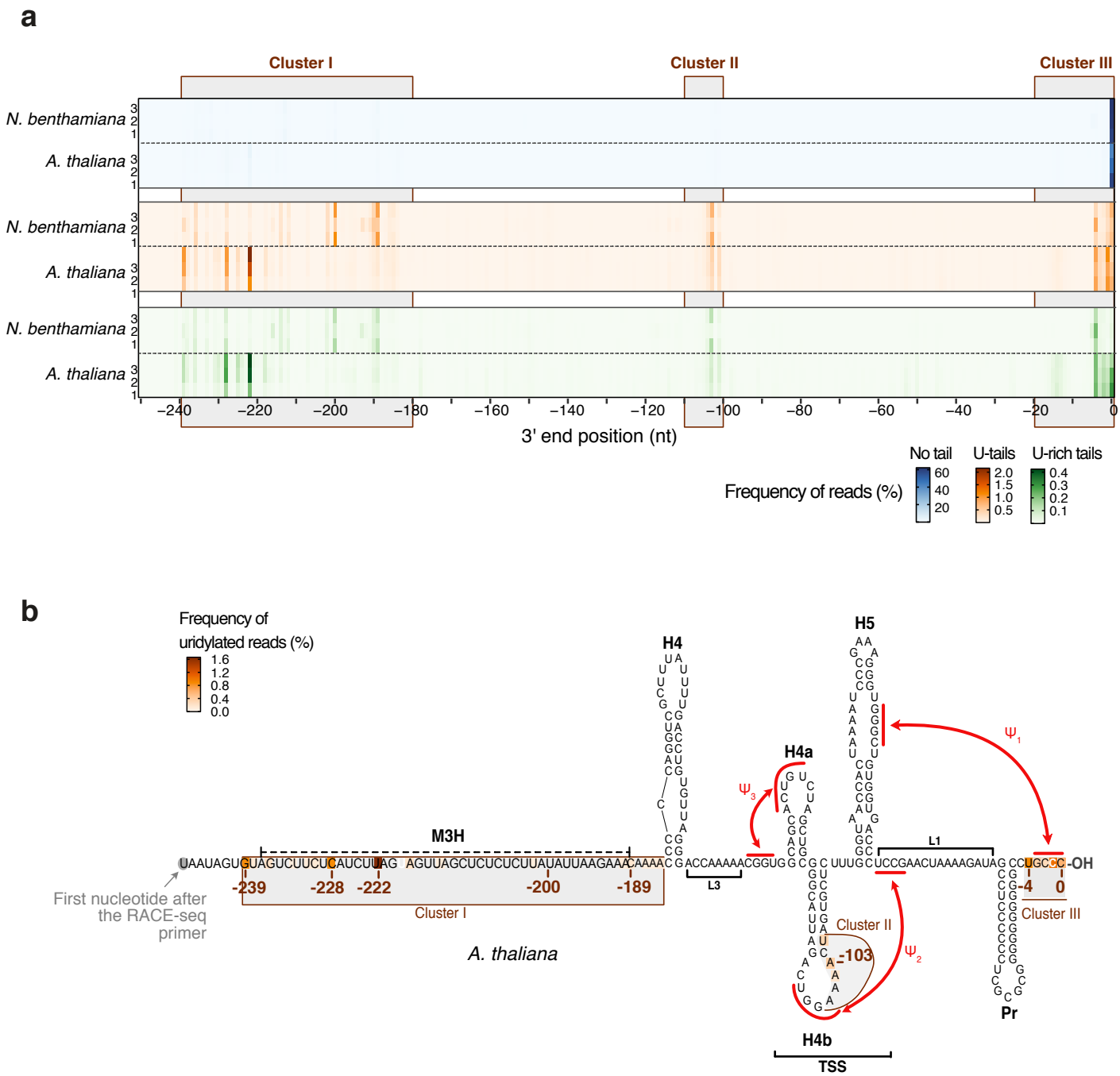


Figure S3: Comparison of TCV RNA uridylation sites detected in infected *N. benthamiana* and Arabidopsis plants. a High resolution mapping of TCV RNA 3' ends in infected *N. benthamiana* and in wild type Arabidopsis plants. Frequencies of reads at each 3' end position are shown by a blue color scale for non-tailed reads, an orange color scale for U-tailed reads and a green color scale for U-rich tailed reads. Position 0 corresponds to the known 3' end. For each host, three virus-infected plants were analyzed. **b** Uridylation sites of TCV RNA in Arabidopsis. See Fig. 3 for the legend and the data obtained in *N. benthamiana.*

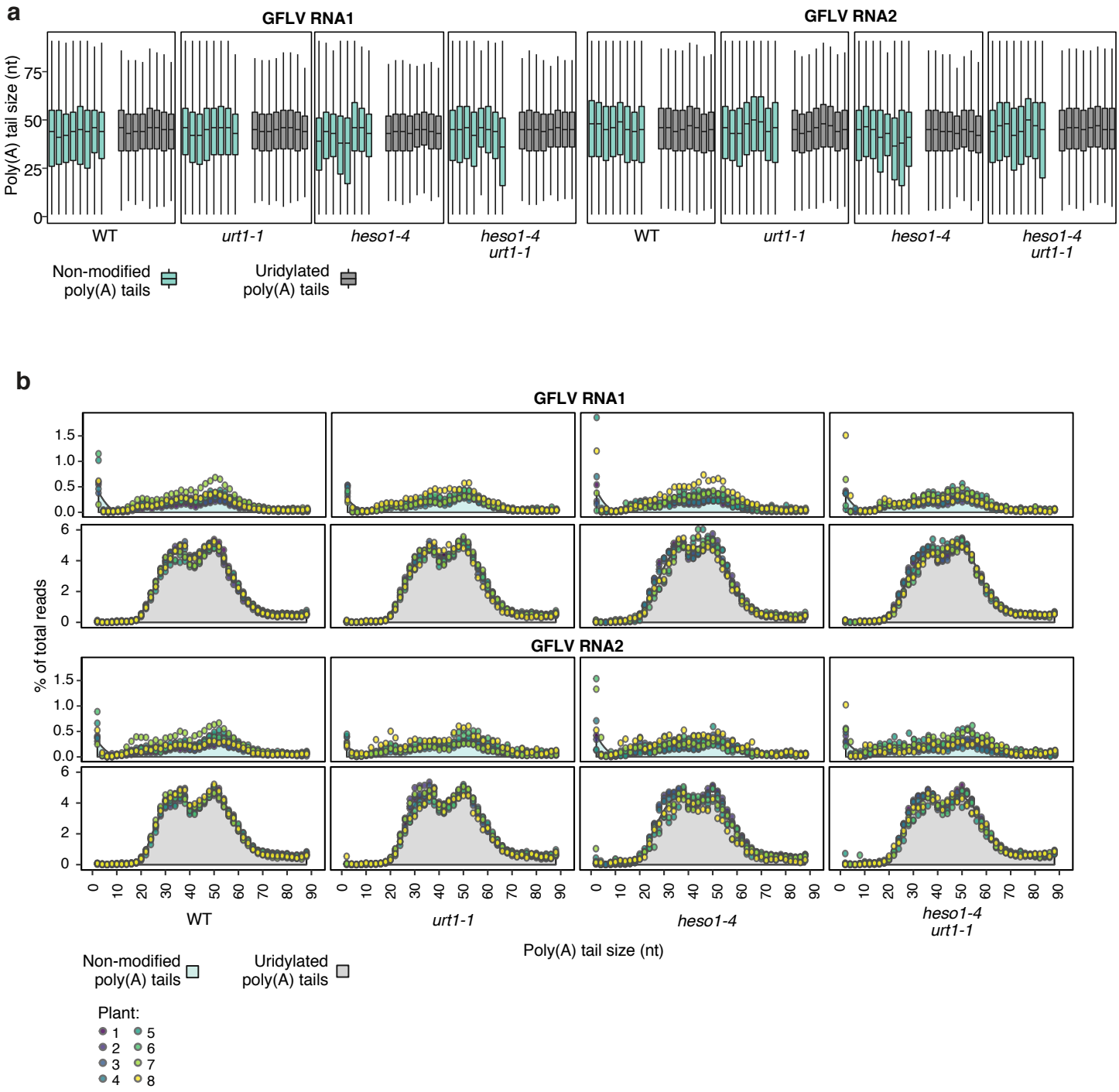


Figure S4 Poly(A) tail size analysis of GFLV RNAs in infected WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* plants. a Boxplot analysis comparing the size of non-modified poly(A) tails (turquoise) vs uridylated poly(A) tails (gray) for GFLV RNAs (GT isolate). Each boxplot represents an infected Arabidopsis plant (n=8) and displays the median, first and third quartile (lower and upper hinges), the largest value within 1.5 times the interquartile range above the upper hinge (upper whisker) and the smallest value within 1.5 times the interquartile range below the lower hinge (lower whiskers). Only viral RNAs with at least 50 U-tailed reads for each replicate are shown. **b** Distribution of poly(A) tail sizes for non-tailed (turquoise) or uridylated (gray) GFLV RNAs (GT isolate) in infected WT, *urt1-1*, *heso1-4* and *heso1-4 urt1-1* plants (Arabidopsis). Percentages were calculated using the total number of sequences with tails from 1 to 89 nucleotides as denominator. Individual points are color-coded for each replicate (n=8) and the average of all replicates is indicated as a gray area.

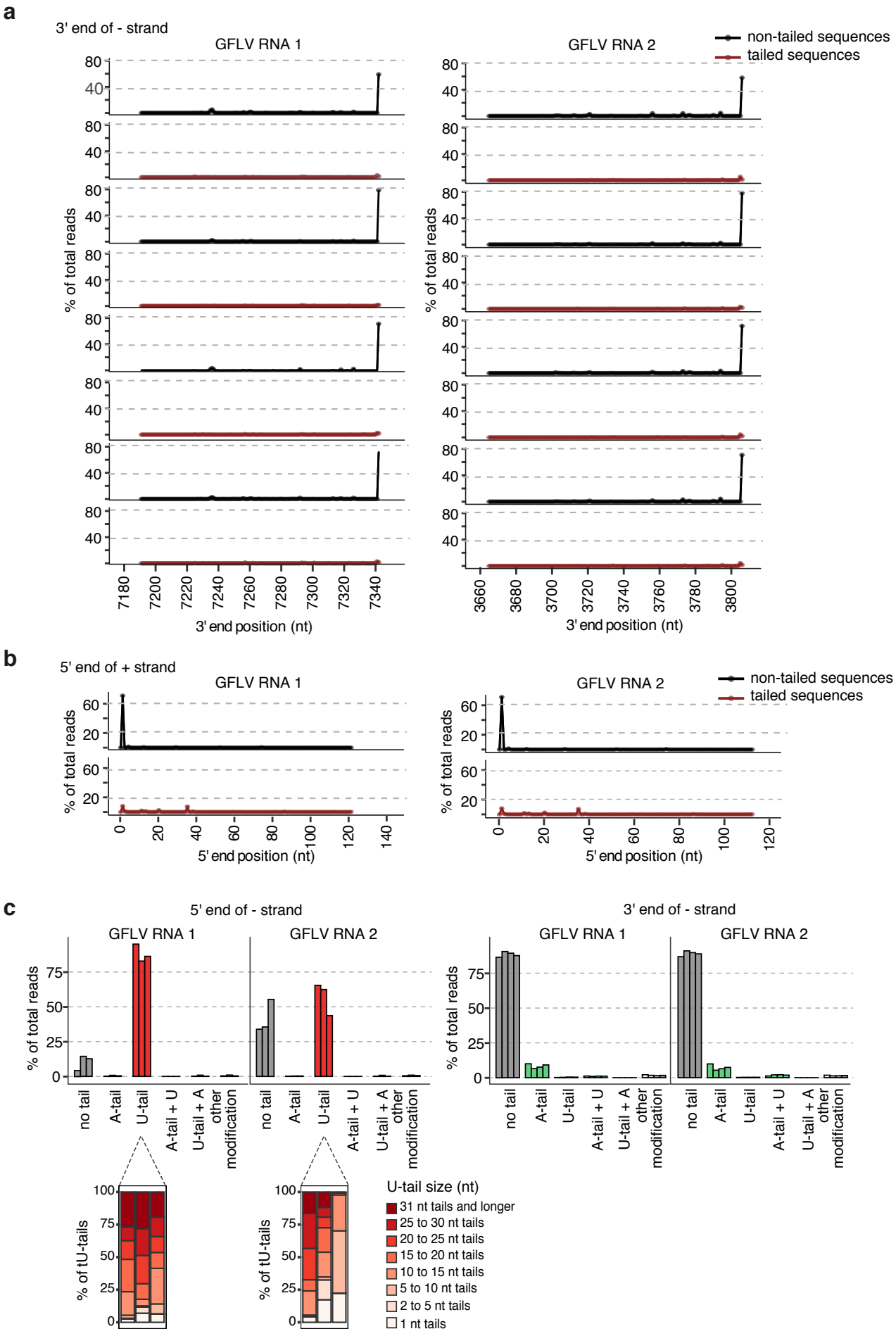


Figure S5 U-tails decorate the 5' extremity of the GFLV RNA minus strand. **a,b** Mapping of the minus strand 3' extremity (a) and the plus strand 5' extremity (b) for GFLV RNAs in infected Arabidopsis plants. Mapping profiles are shown for non-tailed (black) and tailed (red) sequences. **c** Percentages of nucleotide additions at the 5' and 3' end of the minus stand. Proportions of the different tail sizes are shown for U-tails detected at the 5' end of the minus strand. The percentages were calculated using the number of U-tails as denominator. U-tail sizes are indicated by a red gradient from light red, for 1-nt U-tails, to dark red, for 31-nt U-tails and longer. Each bar represents an infected plant (n=3 for 5' end and n=4 for 3' end).

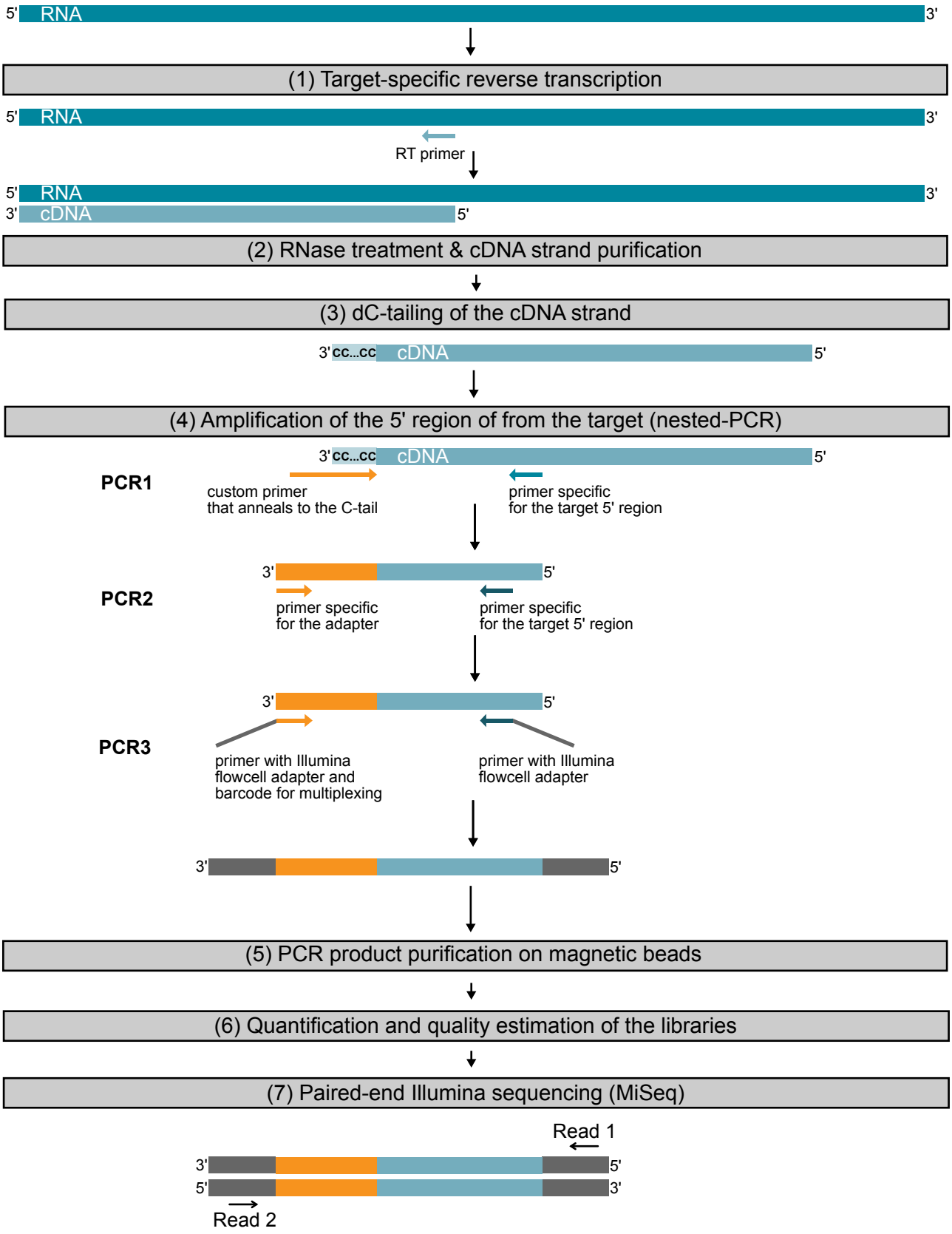


Figure S6 Overview of the 5'RACE-seq workflow. (1) Using total RNA as template, a specific cDNA strand is synthetized using a target-specific primer. (2) The RNA is degraded by an RNase cocktail and the cDNA strand is purified. (3) The 3' end of the cDNA strand is tailed with deoxycytosines: this homopolymeric tail allows the subsequent binding of the PCR1 primer. (4) Three rounds of PCR amplify a product corresponding to the 5' region or the target RNA. In PCR1, one primer is specific to the target. The second primer contains a GI-anchor that anneals to the C-tail of the cDNA, a delimiter sequence, a random sequence of 8 nucleotides and a sequence that will allow the binding of the PCR2 primer. PCR2 and PCR3 primers anneals to the target and the 3' anchor region. PCR3 primers contain Illumina flowcell adapter sequences as well a barcode, which allows a multiplexed analysis. (5) PCR products are purified using magnetic beads. (6) The libraries are quantified and their size is estimated on a Bioanalyzer. (7) The libraries are paired-end sequenced using a MiSeq device: Read 1 allows the identification of the target. Read 2 provides the information about the 5' extremity of the target.

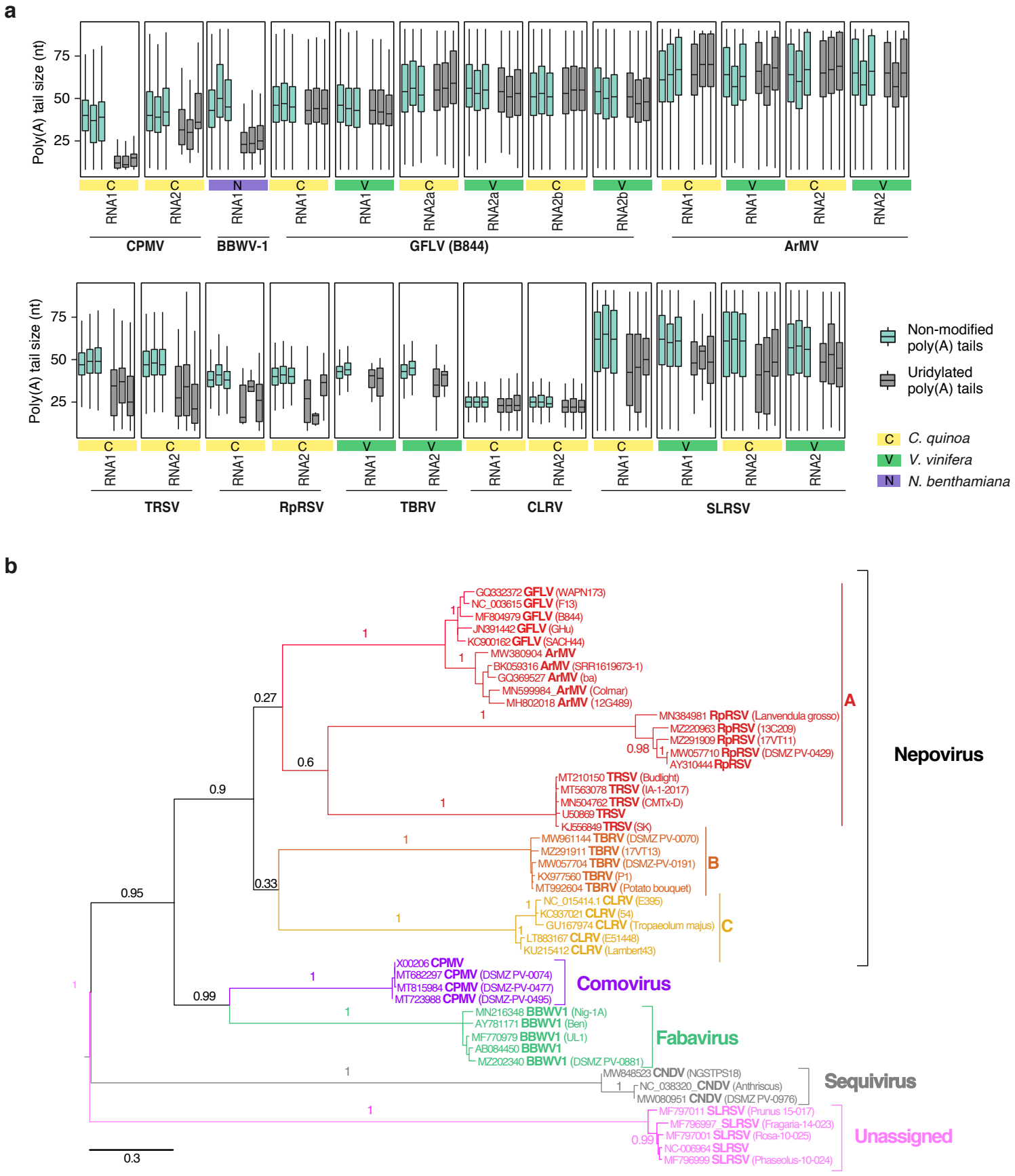


Figure S7 Analysis of RNA 3' ends for different members of the *Secoviridae* family a Boxplot analysis comparing the size of non-modified poly(A) tails (turquoise) vs uridylated poly(A) tails (gray) for plants infected with the indicated viruses. Each boxplot represents an infected plant of the indicated host (n=3) and displays the median, first and third quartiles (lower and upper hinges), the largest value within 1.5 times the interquartile range above the upper hinge (upper whisker) and the smallest value within 1.5 times the interquartile range below the lower hinge (lower whiskers). Only viral RNAs with at least 50 U-tailed reads for each replicate are shown. **b** Phylogenetic tree showing relationships among the species of the *Secoviridae* family analyzed in this study. The tree was calculated from the aligned amino acid sequences of the conserved protease-polymerase (Pro-Pol) region, from the protease CG motif to the polymerase GDD motif. The tree is drawn to scale with branch lengths indicating the number of substitutions per site. Bootstrap values (1000 replicates) are shown above the branches.

Virus name		Order	Family	Subfamily	Genus	5' feature	3' feature	gRNA
Turnip yellow mosaic virus	TYMV	*Tymovirales*	*Tymoviridae*	-	*Tymovirus*	Cap	TLS	1
Peanut clump virus	PCV	*Martellivirales*	*Virgaviridae*	-	*Pecluvirus*	Cap	TLS	2
Tobacco mosaic virus	TMV	*Martellivirales*	*Virgaviridae*	-	*Tobamovirus*	Cap	TLS	1
Tobacco rattle virus	TRV	*Martellivirales*	*Virgaviridae*	-	*Tobravirus*	Cap	No TLS	2
Cucumber mosaic virus	CMV	*Martellivirales*	*Bromoviridae*	-	*Cucumovirus*	Cap	TLS	3
Alfalfa mosaic virus	AlMV	*Martellivirales*	*Bromoviridae*	-	*Alfamovirus*	Cap	No TLS	3
Turnip yellows virus	TuYV	*Sobelivirales*	*Solemoviridae*	-	*Polerovirus*	VPg	No TLS	1
Sowbane mosaic virus	SoMV	*Sobelivirales*	*Solemoviridae*	-	*Sobemovirus*	VPg	No TLS	1
Turnip crinkle virus	TCV	*Tolivirales*	*Tombusviridae*	-	*Betacarmovirus*	-	No TLS	1
Tomato bushy stunt virus	TBSV	*Tolivirales*	*Tombusviridae*	-	*Tombusvirus*	-	No TLS	1
Grapevine leafroll-associated virus 1	GLRaV-1	*Martellivirales*	*Closteroviridae*	-	*Ampelovirus*	Cap	No TLS	1
Carrot necrotic dieback virus	CNDV	*Picornavirales*	*Secoviridae*	-	*Sequivirus*	VPg	Unknown	1
Grapevine leafroll-associated virus 2	GLRaV-2	*Martellivirales*	*Closteroviridae*	-	*Closterovirus*	Cap	A-rich	1
Turnip mosaic virus	TuMV	*Patatavirales*	*Potyviridae*	-	*Potyvirus*	VPg	polyA	1
Beet necrotic yellow vein virus	BNYVV	*Hepelivirales*	*Benyviridae*	-	*Benyvirus*	Cap	polyA	4
Potato virus X	PVX	*Tymovirales*	*Alphaflexviridae*	-	*Potexvirus*	Cap	polyA	1
Grapevine rupestris stem pitting-associated virus	GRSPaV	*Tymovirales*	*Betaflexviridae*	*Quinvirinae*	*Foveavirus*	Cap	polyA	1
Grapevine Pinot gris virus	GPGV	*Tymovirales*	*Betaflexviridae*	*Trivirinae*	*Trichovirus*	-	polyA	1
Grapevine virus B	GVB	*Tymovirales*	*Betaflexviridae*	*Trivirinae*	*Vitivirus*	Cap	polyA	1
Grapevine redglobe virus	GRGV	*Tymovirales*	*Tymoviridae*	-	*Maculavirus*	Cap	polyA	1
Grapevine fanleaf virus	GFLV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus*	VPg	polyA	2

Table 1: Representative species of positive-strand RNA phytoviruses analyzed in this study classified by 3' extremity type and classification (from ICTV Master Species List 2021.v1). The 5' and 3' features are listed according their description in literature and this study. VPg: viral protein genome-linked, TLS: tRNA-like structure. The gRNA column indicates the number of genomic RNA for each virus.

Virus name		Order	Family	Subfamily	Genus	5' feature	3' feature	gRNA
Cowpea mosaic virus	CPMV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Comovirus*	VPg	polyA	2
Broad bean wilt virus 1	BBWV-1	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Fabavirus*	VPg	polyA	2
Raspberry ringspot virus	RpRSV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus A*	VPg	polyA	2
Tobacco ringspot virus	TRSV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus A*	VPg	polyA	2
Grapevine fanleaf virus	GFLV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus A*	VPg	polyA	2
Arabis mosaic virus	ArMV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus A*	VPg	polyA	2
Tomato black ring virus	TBRV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus B*	VPg	polyA	2
Cherry leaf roll virus	CLRV	*Picornavirales*	*Secoviridae*	*Comovirinae*	*Nepovirus C*	VPg	polyA	2
Strawberry latent ringspot virus	SLRSV	*Picornavirales*	*Secoviridae*	-	*Unassigned*	VPg	polyA	2
Carrot necrotic dieback virus	CNDV	*Picornavirales*	*Secoviridae*	-	*Sequivirus*	VPg	unknown	1

Table 2: Representative species of *Secoviridae* analyzed in this study classified by 3' extremity type and classification (from ICTV Master Species List 2021.v1). The 5' and 3' features are listed according their description in literature and this study. VPg: viral protein genome-linked, TLS: tRNA-like structure. The gRNA column indicates the number of genomic RNA for each virus.

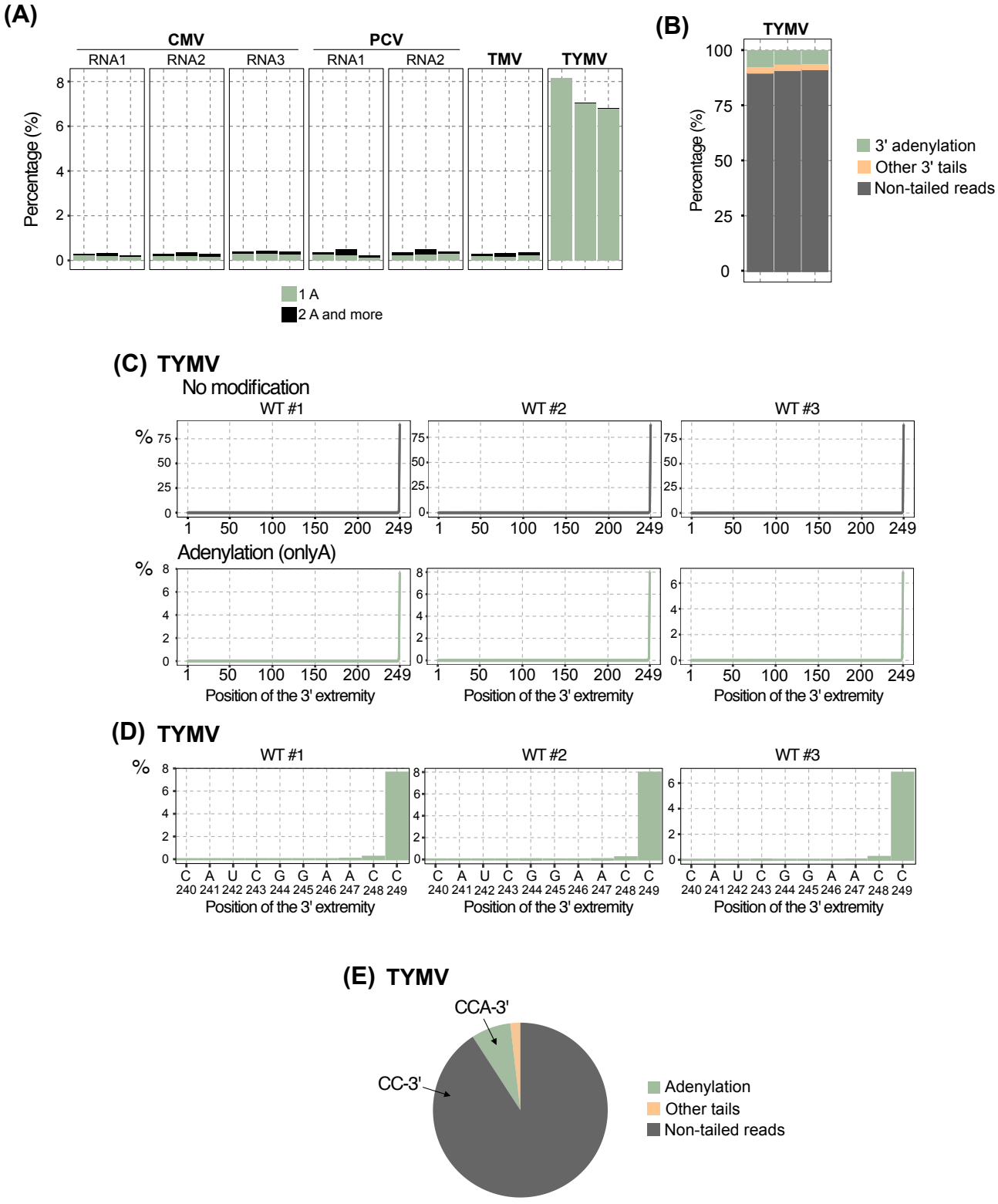


Figure 16: *In vivo* 3' adenylation of TLS-ending viral RNAs. (A) Percentage of adenylation on CMV, PCV, TMV and TYMV RNAs in three different plants. **(B)** Percentage of non-tailed (grey), 3' adenylated (green) and other types of tailed reads (orange) for TYMV RNA in three different plants. Percentages were calculated using the total number of reads for each RNA by plant as denominator. **(C)** High resolution mapping of TYMV RNA 3' extremities in wild-type plants. Frequencies of reads at each 3' end position are shown for non-tailed reads (grey) and A-tailed reads (green). The position 249 corresponds to the expected 3' extremity. **(D)** High resolution mapping of TYMV RNA 3' extremities in wild-type plants. Frequencies of reads on the last 10 3' end positions are shown for A-tailed reads. The nucleotide and the corresponding position are indicated below the graph. The frequencies have been calculated using the number of reads for TYMV RNA by plant as denominator. **(E)** Pie chart recapitulating TYMV RNA 3' extremity.

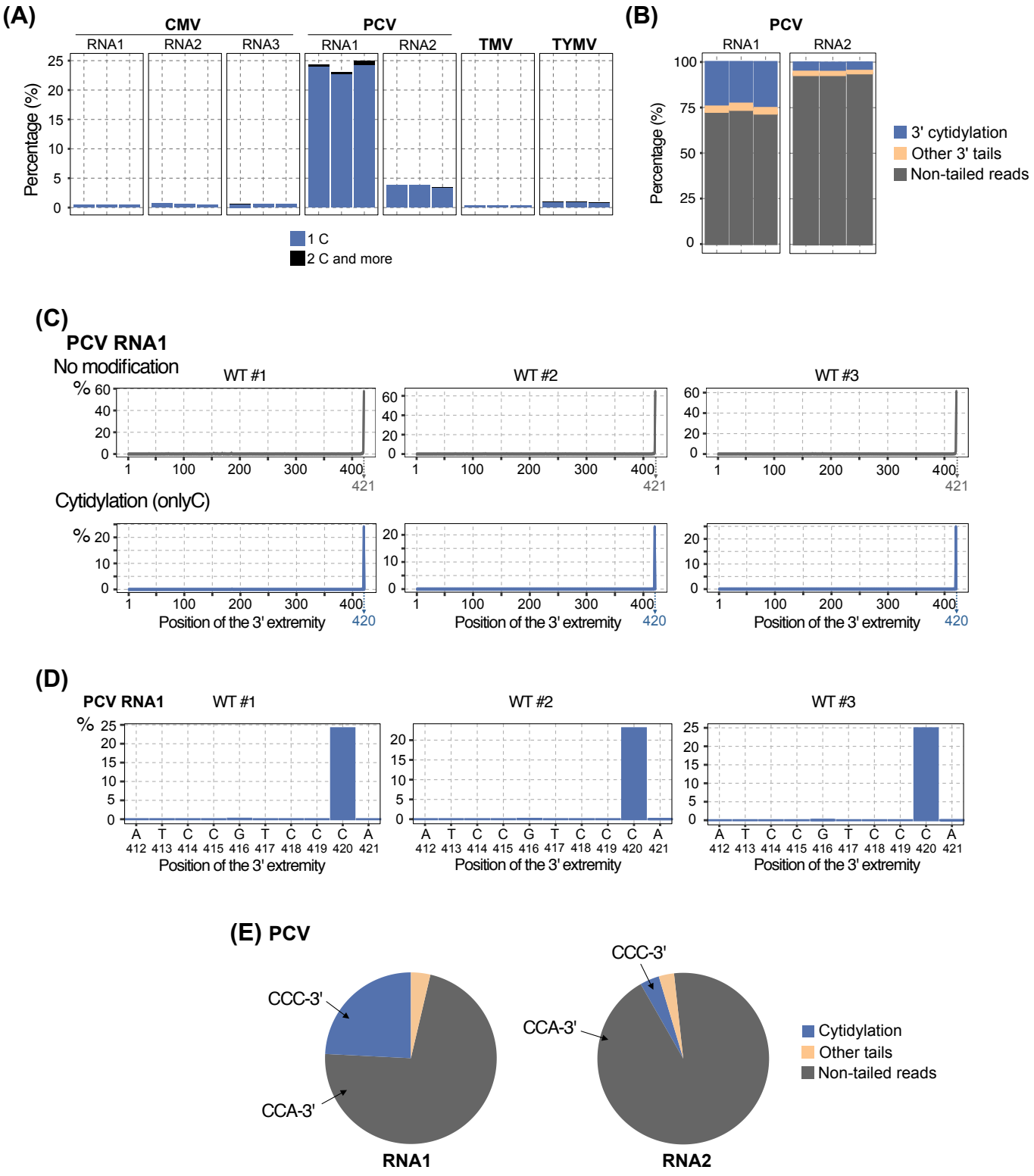
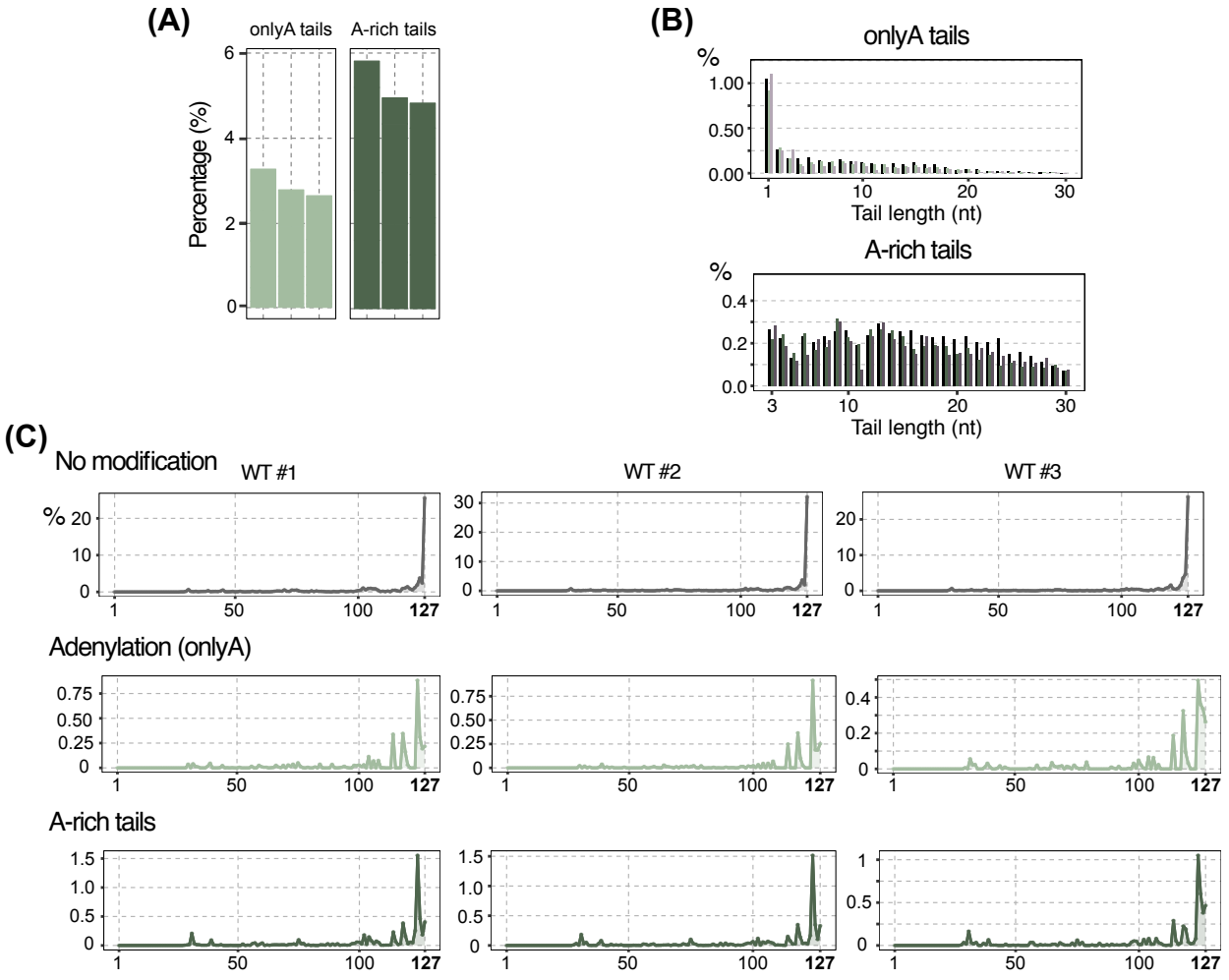


Figure 17: *In vivo* 3' cytidylation of TLS-ending viral RNAs. (A) Percentage of cytidylation on CMV, PCV, TMV and TYMV RNAs in three different plants, calculated using the total number of reads for each RNA by plant as denominator. **(B)** Percentage of non-tailed (grey) and tailed reads (orange) for PCV RNAs in three different plants, calculated using the total number of reads for each RNA by plant as denominator. **(C)** High resolution mapping of PCV RNA1 3' extremities in wild-type plants. Frequencies of reads at each 3' end position are shown for non-tailed reads (grey) and C-tailed reads (blue). Grey and blue arrows show the major peak for non-tailed reads (grey) and cytidylated reads (blue), respectively. The position 421 corresponds to the expected 3' extremity. **(D)** High resolution mapping of PCV RNA1 3' extremities in wild-type plants. Frequencies of reads on the last 10 3' end positions are shown for C-tailed reads. The frequencies in (C) and (D) have been calculated with the number of reads for PCV RNA1 by plant as denominator. The nucleotide and the corresponding position are indicated below the graph. **(E)** Pie charts recapitulating PCV RNAs 3' extremity.



(D)

Examples of A-rich tails detected on GLRaV-1 RNA

```
TTTTAAAAATTTAAAAAAAAATTAAAAAAA
TAAAAAAACAAAGAAAAGATATAAAATTT
AAGAATTAAAAAAAAAAAAAAAAAAA
TTTTAAAATTTAAAAAAAAAAAAAAG
TAATTTAAATAAAAAAAAAAAAAA
AATTAAAAAAAAAAAAAAGAAAAAG
AAAAAAAAAAAAAAAAAATTATT
AAAAAATTAAATTAAAAAAAATT
TTATTTAAAAAAAAAAAAAAA
AAATTAAAACAAAAAAAAAAA
TTTAAAAAAAAAGAGAAAAATT
AATTTTTAAAAAAAAAAAAG
AAATAAAAATAAAAAAAAAG
AAAACAAAAAAAAAAAAAAC
AAAAAAAAATTAAAAATT
TTAAATAAAAAAAAAAAA
TTTTTAAAAAAAAAA
AAAAAAAAAAAAAAC
AAAAAAAAAAAAAT
TCAAAAAAAAAATT
AATTTAAAAAAA
AAAAAAAAATTT
```

Figure 18: GLRaV-1 RNA is adenylated at its 3' extremity (A) Barplot showing the percentage of adenylated reads (onlyA and A-rich tails) for GLRaV-1 RNA. **(B)** Barplot showing the distribution of the length of the A- and A-rich tails detected on GLRaV-1 RNA. On each barplot, each bar represents one plant. **(C)** High resolution mapping of GLRaV-1 3' extremities in wild-type plants. Frequencies of reads at each 3' end position are shown for non-tailed reads (grey), onlyA tails (light green) and A-rich tails (dark green). The percentage in (A) and the frequencies in (B) and (C) have been calculated using the number of reads for GLRaV-1 RNA by plant as denominator. The position 127 corresponds to the last nucleotide of the mature RNA. **(D)** Examples of A-rich tails detected on GLRaV-1 RNA.

3. Other 3' terminal modifications

a. Adenylation and cytidylation of viral RNAs with a 3' TLS

3'RACE-seq allows to map the 3' extremities of target RNAs at high resolution and to analyze 3' non-templated nucleotides. Of note, the RNA must end with a 3' terminal hydroxyl group to be analyzed by 3'RACE-seq. TLS of viral RNAs can be adenylated and aminoacylated *in vitro* (Dreher, 2010). Aminoacylation would prevent the detection by 3'RACE-seq. Hence, in our data, we captured non-aminoacylated TLS.

Our 3'RACE-seq data *in planta* show that the levels of 3' terminal adenylation are very low for CMV, PCV and TMV RNAs (< 0.5%), whereas 6-8% of TYMV RNAs are adenylated at their 3' extremity (Fig. 16A). In fact, terminal adenylation is the major modification that we detected on TYMV RNA (Fig. 16B). Our data show that we efficiently map the expected 3' extremity of TYMV RNA for non-tailed reads (Fig. 16C). The mono-adenylation is detected at the expected 3' extremity of TYMV RNA (Fig. 16C). Moreover, when we look closer at the exact 3' end position modified with a non-templated adenosine, we noticed that almost all of them are added to the last nucleotide of TYMV RNA (Fig. 16D). Overall, our data indicate that most TYMV RNA end in CC_{OH}-3' (90%) while there is a non negligeable proportion (7%) that is adenylated, resulting in CCA_{OH}-3' (Fig. 16E).

Our data show another 3' terminal modification on a TLS, cytidylation. For PCV, we detected almost 25% cytidylation for RNA1 and 4% RNA2, which is, like terminal adenylation, almost exclusively a single cytosine (Fig. 17A). Most PCV RNAs are not tailed, but noteworthy, the proportion of PCV RNA1 that is tailed is much higher than for RNA2 (Fig. 17A, B). We observed that the majority of the reads for PCV RNA1 that are not modified with non-templated nucleotides map to the expected mature extremity, indicated by the arrow (Fig. 17C). The cytidine addition is detected on the penultimate nucleotide (Fig. 17C, D). Overall, for PCV RNAs, our results reveal that whereas most of PCV RNA1 and RNA2 end in CCA_{OH}-3' (Fig. 17E), a high proportion of PCV RNAs ends in CCC_{OH}-3', especially for RNA1 (24%) and at a lesser extent for RNA2 (4%) (Fig. 17E).

b. Detection of A-rich tails on GLRaV-1 RNA

GLRaV-1 RNA, from one of the two viruses of the *Closteroviridae* family that we analyzed, is also adenylated, but to a lesser extent than GLRaV-2 RNA (Fig. 18A, Article Fig 1). In total, about 10% of GLRaV-1 RNAs are decorated with non-templated adenosines, 3% of the tails containing exclusively adenosines and 5-6% of the tails containing a majority of adenosines mixed with other nucleotides (Fig. 18A). Interestingly, the extensions are quite long, especially for A-rich tails that display a relatively widespread distribution up to 30 nucleotides (Fig. 18B). The mapping of GLRaV-1 RNA 3' extremities reveals that reads with no added nucleotides map to the expected mature 3' extremity at position 127. However, the adenylated extensions are detected on 3' termini of nibbled RNAs (Fig. 18C). It is interesting to note that other nucleotides contained in the A-rich tails are mostly uridines, corresponding to the Ts in the examples because we sequenced DNA (Fig. 18D).

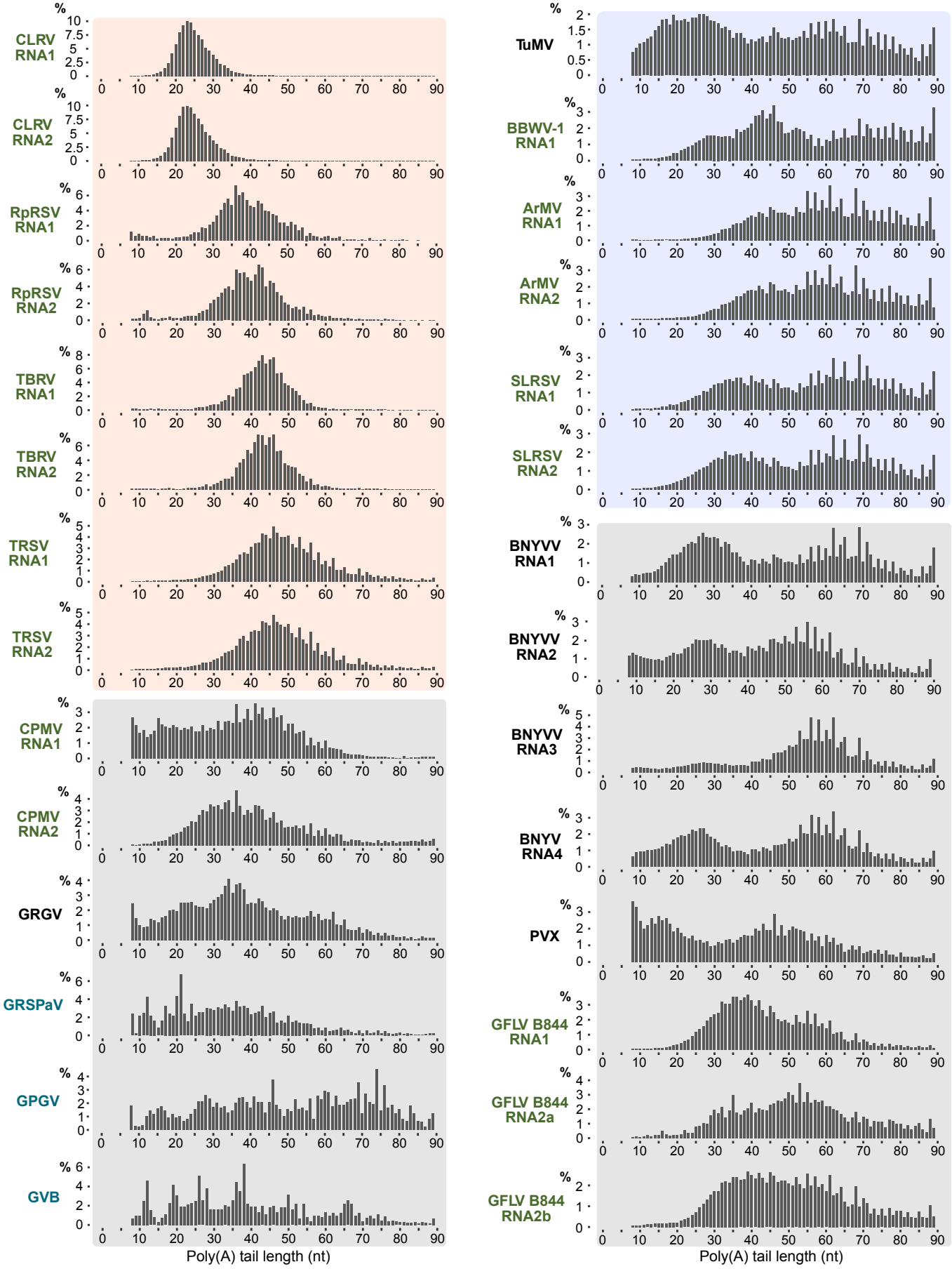


Figure 19: The distribution of the poly(A) tails profiles on plant viral RNAs is diverse. The frequency of reads by poly(A) tail size is plotted for viral RNAs that have more than 500 polyadenylated reads, combining the data of three independent plants (GVB two plants). The target was identified in the read 1 and in associated read 2, a poly(A) tail of at least 8 adenosines was searched. Due to the analysis pipeline limitations, the frequencies of the tails from 8 to 89 As are plotted. The frequencies have been calculated with the number of reads for each RNA as denominator. Species from the *Secoviridae* and *Betaflexviridae* are written in green and in blue, respectively. The colored rectangles (light orange, purple and grey) indicate the three groups according to the poly(A) tail size distribution.

4. Diversity of the poly(A) tail profiles between viral RNAs

3'RACE-seq data provide information on the size and the distribution of the poly(A) tail. One of the main interests of our study is the resolution of the poly(A) tail profiles of plant viral RNAs at a high resolution *in vivo* for 16 plant viral species. Fig. 19 shows poly(A) tail distribution profiles for 16 viruses, including nine species from the *Secoviridae*. Of note, for some of the viral RNAs of the *Secoviridae*, we do not have a complete reference sequence, and thus were not able to map RNA 3' position and to analyze poly(A) tails shorter than 8 As (see Material and Methods). So, I plotted the poly(A) tail distribution from 8 to 89 As. Also, in order to obtain reliable distribution, the poly(A) tail size distribution was only plotted for viral RNAs with at least 500 reads in the final dataset, thereby excluding BBWV-1 RNA2.

The first obvious observation is the diversity of the distribution of the poly(A) tail sizes observed for the different viral RNAs. We can roughly divide viral RNAs in three groups according to their poly(A) tail size distribution profile (Fig. 19). The first group contains viral RNAs that show a narrow unimodal distribution of poly(A) tail sizes: CLRV, RpRSV, TBRV, and TRSV RNAs. Among them, CLRV RNAs have much shorter tails compared to others. Its distribution ranges from 15 to 35 As while distributions of RpRSV and TBRV RNAs ranges from 25 to 60 As.

The second group of viral RNAs shows broad distribution profiles of poly(A) tail sizes with a high proportion of poly(A) > 70 As. This group includes TuMV, BBWV-1, ArMV and SLRSV RNAs. In the case of TuMV, the distribution is relatively uniform from 8 to 89 adenosines.

The last group of viral RNAs display intermediary broad profiles with modest or low amount of poly(A) tails longer than 70 As. Among these viral RNAs, CPMV, GRGV, GRSPaV and PVX RNAs do not present longer tails than 60 As, while GFLV and BNYVV RNAs present non negligible amounts of poly(A) longer than 70As.

Interestingly, many viral RNAs from the two last groups often show bimodal poly(A) tail size distribution. For example, three of the four RNAs of BNYVV display two clear peaks at 25 and ca. 60 As, respectively. It is interesting to note the RNA3 of BNYVV also has a peak at 25 As, although the second peak at 60 As is predominant. Of note, PVX is the viral RNA that accumulates the highest proportion of RNAs with poly(A) tail shorter than 10 As, as compared to other viruses that we analyzed.

Intriguingly, we see that even for viral species from the same family, the distribution of poly(A) tail size varies. We analyzed three species of the *Betaflexviridae*, GPGV, GRSPaV and GVB. While GPGV RNA possesses long poly(A) tails, up to 89 As, that are widely distributed, poly(A) tail sizes for GRSPaV and GVB RNAs do not exceed 60 As. The diversity of the poly(A) tail distribution is even more striking within the *Secoviridae* family. We analyzed nine viruses that are all bipartite, so in total we analyzed 18 different viral RNAs. We observed very diverse distributions of poly(A) tail sizes: CLRV, TBRV and RpRSV RNAs show a narrow unimodal distribution whereas the distribution for other *Secoviridae* viruses is wider.

Our analysis of the poly(A) tail of viral RNAs reveals an extreme diversity of the poly(A) tail size. Globally, our data provide an important resource to study the poly(A) tails of plant viral RNAs which could be of use for further work to investigate the synthesis and the role of this 3' feature for viral RNAs in plants.

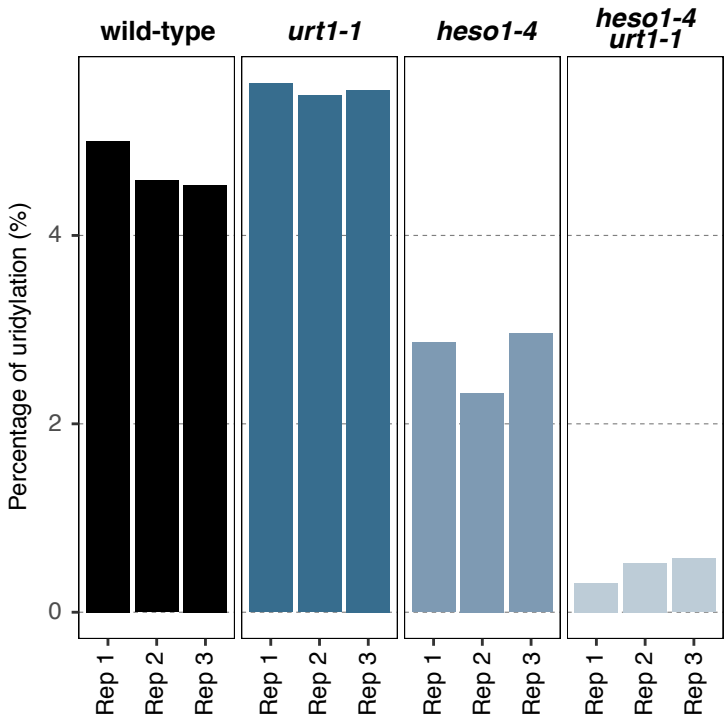


Figure 20: Percentage of uridylation on TuMV RNA at 10 dpi. The percentage of uridylated reads has been calculated with the total number of reads for TuMV RNA in each plant as denominator.

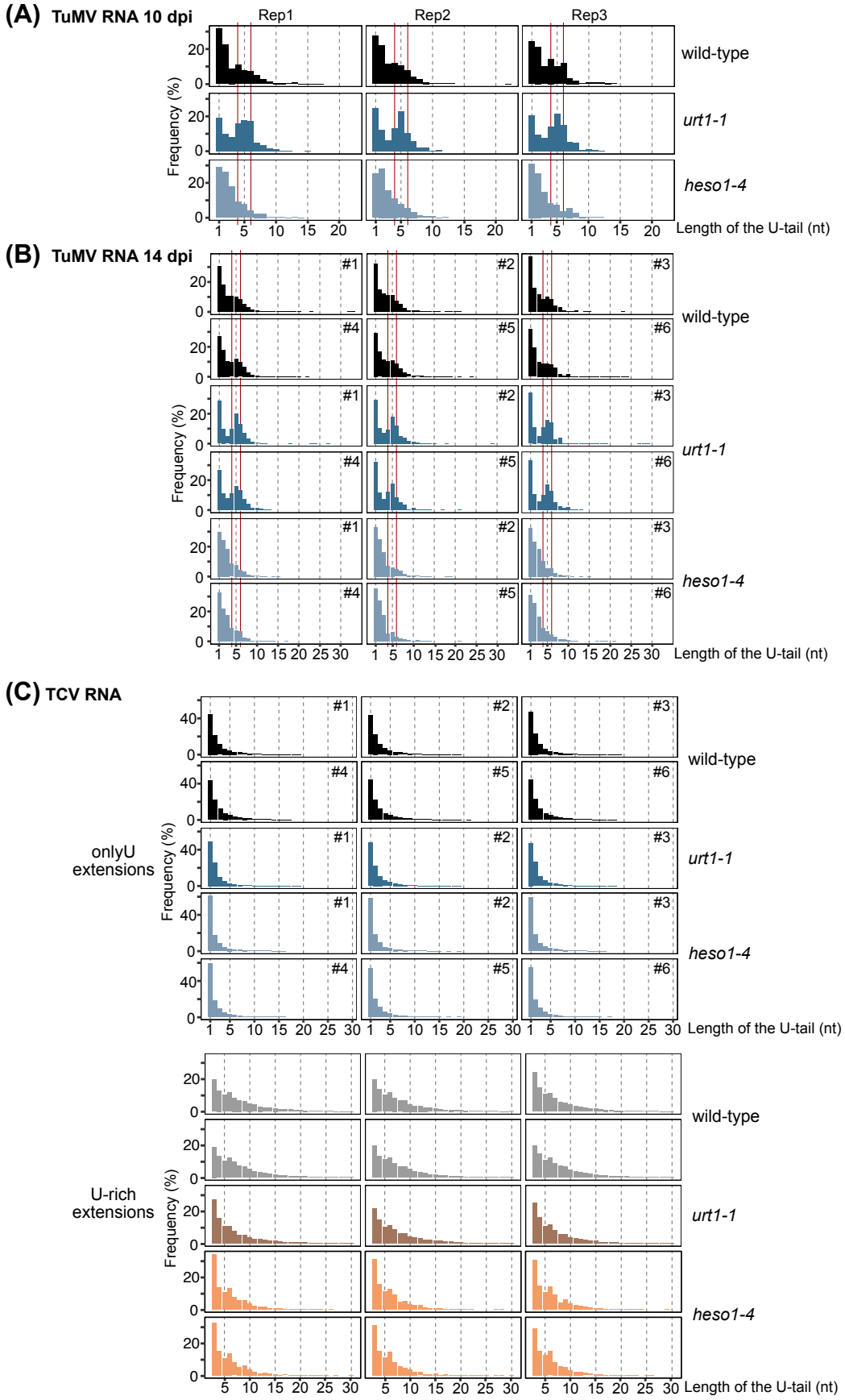


Figure 21: URT1- and HESO1-added extensions have different lengths. Barplots representing the proportion of the reads over the length of the uridine tail for TuMV RNA at **(A)** 10 dpi and **(B)** 14 dpi and **(C)** TCV RNA. For TuMV RNA, the majority of the U-tails represents uridines added to the poly(A) tail. At 10 dpi, 2-3 plants per genotype for 3 biological replicates have been pooled for TuMV RNA uridylation analysis. At 14 dpi, TuMV RNA uridylation has been determined in 3 plants of 2 independent biological replicates (#1-3 and #4-6). TCV RNA uridylation has been studied in 3 plants of 2 independent biological replicates (#1-3 and #4-6) for wild-type and *heso1-4*, and 3 plants of 1 replicate for *urt1-1*. The frequencies have been calculated with the number of uridylated reads (onlyU in black/blue and U-rich tails in grey/orange) for each RNA by plant as denominator.

5. URT1 and HESO1 add extensions of different sizes

From previous studies, we know that URT1 and HESO1 have preferred RNA substrates. URT1 is the main TUTase that uridylates mRNAs (Sement et al., 2013; Zuber et al., 2016) and HESO1 is the main TUTase that targets small RNAs and RISC-cleaved transcripts (Ren et al., 2012, 2014; Zuber et al., 2018). In the case of viruses, our data show that both TUTases can uridylate viral RNAs. TCV non-polyadenylated RNA is uridylated by both TUTases as the inactivation of URT1 and HESO1 impacts the global uridylation level (Article Fig. 4A). TuMV polyadenylated RNA is uridylated by URT1 and, to a lesser extent, by HESO1 (Article Fig. 6A). Yet, we noticed a differential preference of URT1 and HESO1 for the last recognized 3' nucleotide on TCV RNA, which is preferentially an A for URT1 and non-A nucleotides for HESO1 (Article Fig. 4C-E). Interestingly, our results also highlighted another difference in the activity of these TUTases, only for TuMV RNA: both TUTases add uridine extensions of different sizes.

Our data indicate that TuMV RNA is uridylated at 3-5% at 10 and 14 days post inoculation (dpi) and that there is almost no more uridylation in the double mutant *heso1-4 urt1-1* (Article Fig. 6A; Fig. 20). To determine the number of U added by URT1 and HESO1, I studied the length of U extensions for the two single mutants. Our results show an increased proportion of U-tails between 4 and 6 Us in *urt1-1* (40%), as compared to *heso1-4* (20%) and also to wild-type (28%), but to a lesser extent (Fig. 21A, B). So, our data indicate that, while both URT1 and HESO1 uridylate TuMV RNA, URT1 adds shorter extensions than HESO1.

In wild-type plants, ca. 10% and 3% of TCV RNA is uridylated with oligo(U) or U-rich tails, respectively (Article Fig. 4A). The absence of both TUTases results in uridylation at background level (Article Fig. 4A). As for TuMV, I investigated the length of the extensions in wild-type and in single mutants. Most U-tails detected on TCV RNAs range from 1 to 4 uridines in wild-type and in the TUTases mutants, while U-rich tails are between 3 and 10 nucleotides (Fig. 21C). Although the proportion of U-rich tails detected on TCV RNA is strongly reduced in *heso1-4* (Article Fig. 4A), it does not impact the length of the U-rich extensions (Fig. 21C). By contrast to TuMV RNA, we did not detect a clear difference in the length of the uridine extensions for TCV RNA.

Our 3'RACE-seq results reveal that URT1 and HESO1 display distinct processivities on viral RNAs. For TuMV polyadenylated RNA, URT1 adds shorter extensions than HESO1. For TCV non-polyadenylated RNA, our data does not reveal a clear difference of processivity between HESO1 and URT1 for tails exclusively composed of uridines. Yet, we show that HESO1 adds long U-rich tails on TCV RNA, which supports the idea that HESO1 is more processive than URT1. Moreover, URT1 preferentially uridylates substrates that end with a 3' terminal A, whereas HESO1 preferentially adds uridines to 3' terminal non-A nucleotides for the substrates that we tested. The differences in processivity and substrate preference could imply distinct functions in viral RNA metabolism.

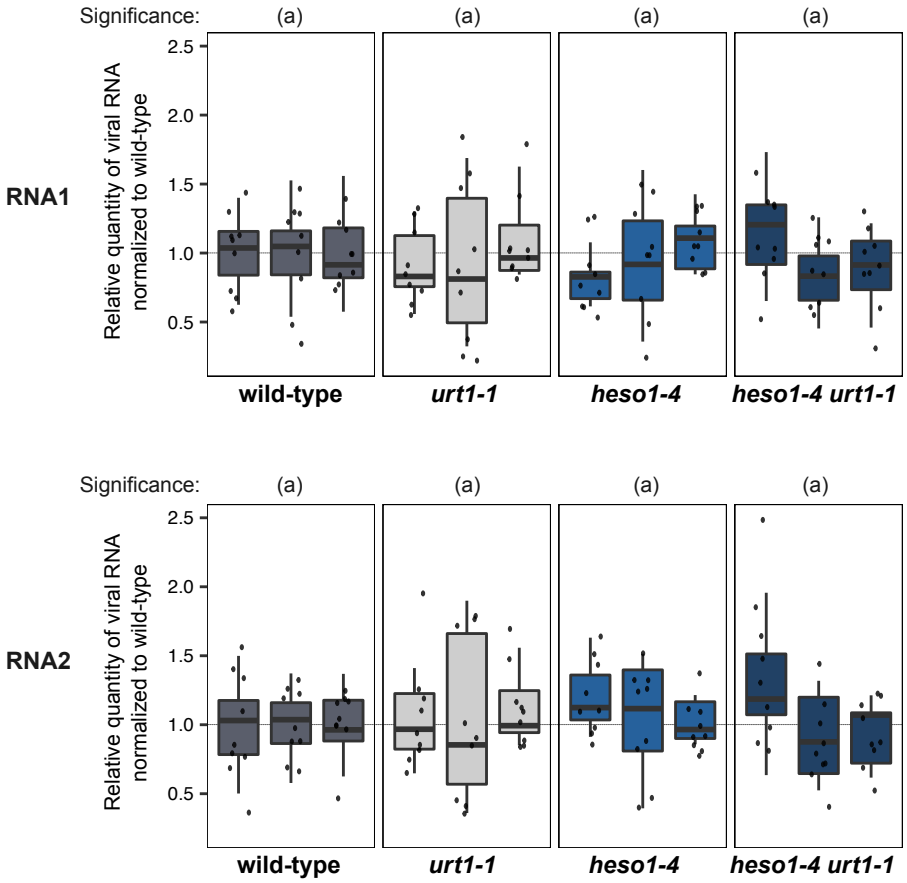


Figure 22: URT1 and HESO1 do not impact GFLV RNAs accumulation at 14 dpi. Boxplots representing the relative quantity of GFLV RNA1 and RNA2 determined by RT-qPCR. The viral RNA quantity has been normalized to Arabidopsis *ACT2* housekeeping gene. The relative quantities have been normalized to the mean relative quantity of the wild-type plants by replicate. Nine infected and one mock plants per genotype for three independent biological replicates have been quantified. Each boxplot represents one biological replicate and black points correspond to individual plants. Letters show significant differences (p-value <0.05). The complete statistical analysis is shown in supplemental table 5.

2. Study of the impact of uridylation on viral accumulation

1. The absence of the TUTases does not affect systemic viral accumulation

To date, there is no data available on the impact of uridylation on the viral RNA accumulation in plants. We detected various levels of uridylation on viral RNAs with singular uridylation patterns, in particular on GFLV RNAs (Article Fig. 2). Interestingly, in *C. elegans* and mammalian TUTases have been reported to be involved in uridylating viral RNAs, thereby inducing a decrease in their accumulation (Le Pen et al., 2018). We thus investigated in Arabidopsis whether the absence of the host TUTases has an impact on viral accumulation *in vivo*. I chose to analyze the accumulation of viral RNA for two model viruses which RNA uridylation patterns and mechanisms are distinct, GFLV and TuMV (Article Fig. 2).

I first quantified viral RNA accumulation of GFLV and TuMV RNAs by RT-qPCR in wild-type and plants mutated for the TUTases URT1 and HESO1 in Arabidopsis. In addition, the TuMV infectious clone contains the *GFP* gene and the GFP signal, which is visible under UV light, can be used to quantify the TuMV infection to verify that the infection has started and to evaluate the viral accumulation.

a. GFLV accumulation does not change in TUTases mutants

Although we have demonstrated that cellular TUTases are not involved in GFLV RNA uridylation (Article Fig. 7A), we cannot rule out the possibility of an indirect effect of host TUTases on viral accumulation. So, I quantified by RT-qPCR the viral RNA accumulation of GFLV (isolate GT) RNAs after 14 days of infection in systemically infected plants for four genotypes: Col-0, *urt1-1*, *heso1-4*, *heso1-4 urt1-1*. Of note, these infected plants were the same that those analyzed by 3'RACE-seq (Article Fig. 7A). For each of the four genotypes, I inoculated nine plants with the virus and one with buffer as a mock control. I repeated this experiment for three independent biological replicates.

Both GFLV RNAs have been quantified by RT-qPCR and their relative level to the Arabidopsis *ACT2* housekeeping gene is presented for each plant (Fig. S1). As expected, we did not detect any viral RNA in the mock plants. For each genotype, all the inoculated plants are infected, at a relatively similar level. We see that RNA2 accumulates more than RNA1. To normalize the differences between independent experiments, I also calculated the ratio of each quantity as compared to the mean quantity of the wild-type plants (Fig. 22). We did not detect any significant difference in the accumulation of GFLV RNAs between the wild-type and the mutant plants and this observation is reproducible between the three replicates (Fig. 22). This observation is consistent with the fact that URT1 nor HESO1 are not involved in the uridylation of GFLV RNAs (Article Fig. 7A).

b. Mutating HESO1 or URT1 does not impact TuMV systemic infection

I evaluated TuMV accumulation in systematically infected plants at 14 dpi in wild-type and TUTases mutants by RT-qPCR, as well as by quantifying the GFP signal. Of note, 14 dpi correspond to the time point previously analyzed by 3' RACE-seq (Article Fig. 6). For two independent biological replicates, I inoculated nine plants with the virus and one with buffer as a mock control per genotype. The level of accumulation of TuMV RNA is presented relative to Arabidopsis *ACT2* housekeeping gene for each plant (Fig. S2A). As for GFLV, I also normalized each quantity as compared to the mean quantity of the

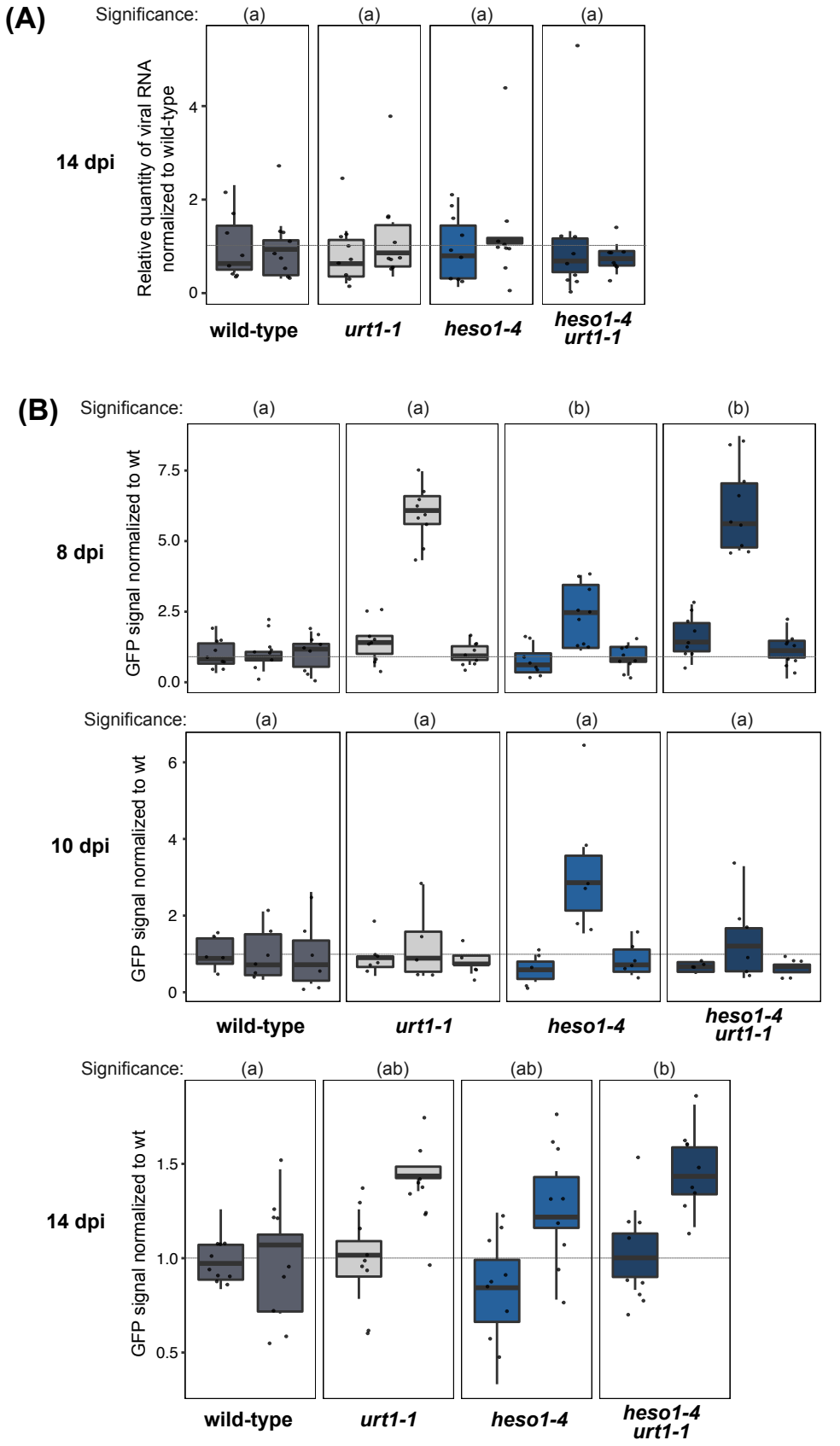


Figure 23: The absence of URT1 and HESO1 does not impact TuMV accumulation. (A) Boxplot representing the relative quantity of TuMV RNA determined by RT-qPCR at 14 dpi. The viral RNA quantity has been normalized to Arabidopsis *ACT2* housekeeping gene. The relative quantities have been normalized to the mean relative quantity of the wild-type plants by replicate. Nine infected and one mock plants per genotype for two independent biological replicates have been quantified. **(B)** Boxplot representing the GFP signal normalized to the mean signal of the wild-type plants by replicate at 8, 10 and 14 dpi. At 8, 10 and 14 dpi, nine, four to six and nine plants per genotype have been quantified, respectively. Each boxplot represents one biological replicate and black points correspond to individual plants. Letters show significant differences (p-value <0.05). The complete statistical analysis is shown in supplemental table 5. The grey line indicates the fold-change of 1, *i.e.*, no change compared to wild-type.

wild-type plants (Fig. 23). We noticed slightly more variability of the relative RNA level between the plants for each genotype, as compared to GFLV and we did not see any significant effect of the absence of URT1 and HESO1 on the accumulation of TuMV RNA at 14 dpi (Fig. 23A). By quantifying the GFP signal, we obtain similar levels in each plant by genotype for both replicates (Fig. S2B). Of note, GFP signal quantification values are reproducible between plants within the same independent replicate but less between two independent replicates, as evidenced by the discrepancy between replicate 1 and 2 (Fig. S2B). Globally, the GFP signal quantification is a little higher in the mutants as compared to the wild-type for the replicate 2 but it is not the case for the replicate 1 (Fig. 23B). So far, our actual data on TuMV accumulation at 14 dpi tend to indicate that the absence of the TUTases does not impact its accumulation.

We have seen that TuMV RNA accumulates at high levels at 14 dpi and that the absence of the TUTases does not affect the level of accumulation. So, I also assessed viral accumulation at earlier timepoints, where the virus accumulates less. For biological replicates 1 and 2, I had photographed the plants at 8 dpi to make sure that the infection started, so these sets of plants are the same than those for replicates 1 and 2 at 14 dpi. I quantified viral accumulation at 8 and 10 dpi for four to nine infected plants per genotype and for three independent replicates. Of note, our 3'RACE-seq data indicate that there is no change in TuMV RNA uridylation pattern at 10 dpi compared to what we observed at 14 dpi: the proportion of uridylated TuMV RNA is ca. 4%. in wild-type plants, decreases in the absence of URT1 and drops to background level in *heso1-4 urt1-1* (Article Fig. 6A; Fig. 20). Both at 8 and 10 dpi, there is an important variability in GFP signal between the plants (Fig. S2B) and it is difficult to detect a reproducible global trend of the viral accumulation in the absence of the TUTases (Fig. 23B). The statistical analysis shows a significant increase of GFP intensity in *heso1-4* (at 8 dpi) and *heso1-4 urt1-1* (at 8 and 14 dpi). Yet, the observed variability, that is particularly important for the second replicate of wild-type plants, could bias statistical analysis (Fig. 23B, Fig. S2B).

Altogether, our analyses did not allow us to highlight an impact of the absence of URT1 and HESO1 on the viral accumulation. Considering the high biological variability observed in our experiments, more replicates and/or an optimization of the experimental protocol are needed to ascertain whether viral RNA accumulation varies in TUTase mutants.

2. TUTase overexpression could negatively affect viral accumulation

a. Generation of transgenic lines overexpressing TUTases

Even though we did not detect major changes in viral accumulation in the absence of URT1 and HESO1, it is not excluded that their overexpression can impact viral accumulation. To investigate this aspect, I used transgenic lines overexpressing tagged versions of URT1 or HESO1. The URT1 overexpressing line has been generated by our team by expressing URT1-GFP under the control of a ubiquitin promoter (*UBI10*) in *urt1-1* plants. During my PhD, I generated different lines expressing HESO1 tagged at its C-terminus with RFP, GFP or 4xmyc, respectively. The expression of the fusion protein is under the control of the ubiquitin or the 35S promoter. All HESO1 constructs were transformed into *heso1-4* background (Table 4).

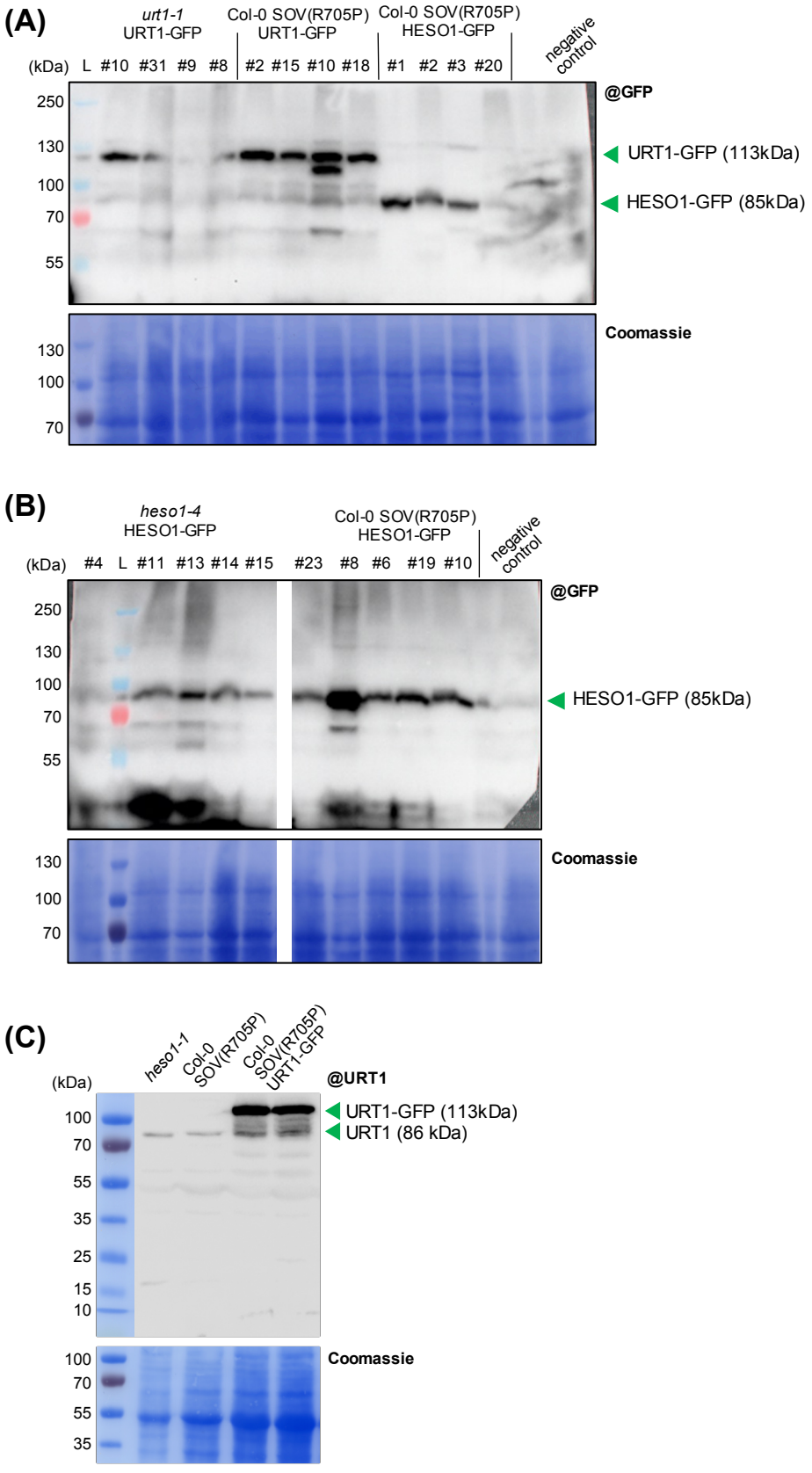


Figure 24: Estimation of URT1-GFP and HESO1-GFP levels of expression in transgenic lines. Western blot analysis in plants expressing URT1-GFP or HESO1-GFP in *urt1-1* or *heso1-4*, respectively, or Col-0 SOV(R705P) backgrounds. Molecular weight markers are indicated in kDa. A part of the membrane stained with Coomassie blue is shown as loading control. **(A) & (B)** Western blot analysis with an anti-GFP antibody. The samples in (B) originate from the same membrane that has been cropped in the middle to remove unrelated samples. Col-0 serves as negative control. **(C)** Western blot analysis with an anti-URT1 antibody. *heso1-1* serves as control for endogenous URT1 expression.

Transgenic line	Plasmid	Promoter	Terminator	Genetic background	Reference
urt1-1 URT1-GFP	pUBC-GFP	Ubiquitin	35S	*urt1-1*	unpublished, H. Scheer and C. De Almeida
heso1-4 HESO1-GFP	pUBC-GFP	Ubiquitin	35S	*heso1-4*	this work
heso1-4 HESO1-RFP	pUBC-GFP	Ubiquitin	35S	*heso1-4*	this work
heso1-4 HESO1-4xmyc	pGWB17	35S	NOS	*heso1-4*	this work
Col-0 SOV (R705P)	pCambia1300	endogenous	endogenous	Col-0	Zhang et al., 2010
Col-0 SOV (R705P) URT1-GFP	pUBC-GFP	Ubiquitin	35S	Col-0 SOV (R705P)	this work
Col-0 SOV (R705P) HESO1-GFP					

Table 4: Transgenic lines generated or used in this study. The name of the transgenic line is indicated in the first column, followed by the plasmid used, the promoter and terminator under which the construct is expressed *in planta*, the genetic background used for the transformation and reference for the line if it is published.

SOV (*AT1G77680*) in plants is the ortholog of DIS3L2, a 3' to 5' exoribonuclease degrading uridylated mRNAs and non-coding RNAs (Łabno et al., 2016; Lubas et al., 2013; Reimão-Pinto et al., 2016; Thomas et al., 2015; Ustianenko et al., 2013, 2016). It is thus reasonable to speculate that SOV in plants could also participate to the degradation of uridylation RNAs, including viral RNAs. However, this has not been yet addressed experimentally. Columbia (Col-0) is the most used accession for studies in Arabidopsis and was used so far for all analyses of this work. In the vast majority of Arabidopsis accessions, SOV contains a proline at residue 705, like the Landsberg (Ler) accession, whereas in Col-0 SOV has an arginine at this position (Zhang et al., 2010). This point mutation in Col-0 could abrogate, or at least impede, the function of SOV in RNA degradation. So, if SOV-mediated decay of uridylated RNAs exists in plants, it does probably not occur in Col-0 accession, that should be considered as a SOV mutant. To assess the impact of a functional version of SOV on uridylation-mediated decay, I also generated plants overexpressing URT1-GFP or HESO1-GFP in a Col-0 SOV (R705P) background (Table 4). The line Col-0 SOV (R705P) has been obtained by transformation of Col-0 plants with SOV that possesses a proline at residue 705, expressed under its endogenous promoter (Zhang et al., 2010). Hence, two versions of SOV can be expressed in this line. Of note, we do not have an antibody against SOV and therefore, we do not know the relative expression level of both versions.

After two or three rounds of selection, I obtained lines expressing HESO1-GFP in *heso1-4*, HESO1-GFP or URT1-GFP in Col-0 SOV (R705P). I performed a western blot analysis on the plants to assess the level of expression in these overexpressing lines, as well in the *urt1-1* URT1-GFP line. As an example, Fig. 24 presents expression patterns for URT1- and HESO1-GFP in the four genetic backgrounds. Most plants express the fusion protein at a similar level (Fig. 24A, B).

We can observe that URT1-GFP is expressed at a similar level between both backgrounds (Fig. 24A, B). In addition, URT1-GFP is overexpressed as compared to endogenous URT1 (Fig. 24C). We tried to produce anti-HESO1 antibodies, unsuccessfully, so, I could not assess HESO1-GFP expression level, as compared to endogenous HESO1.

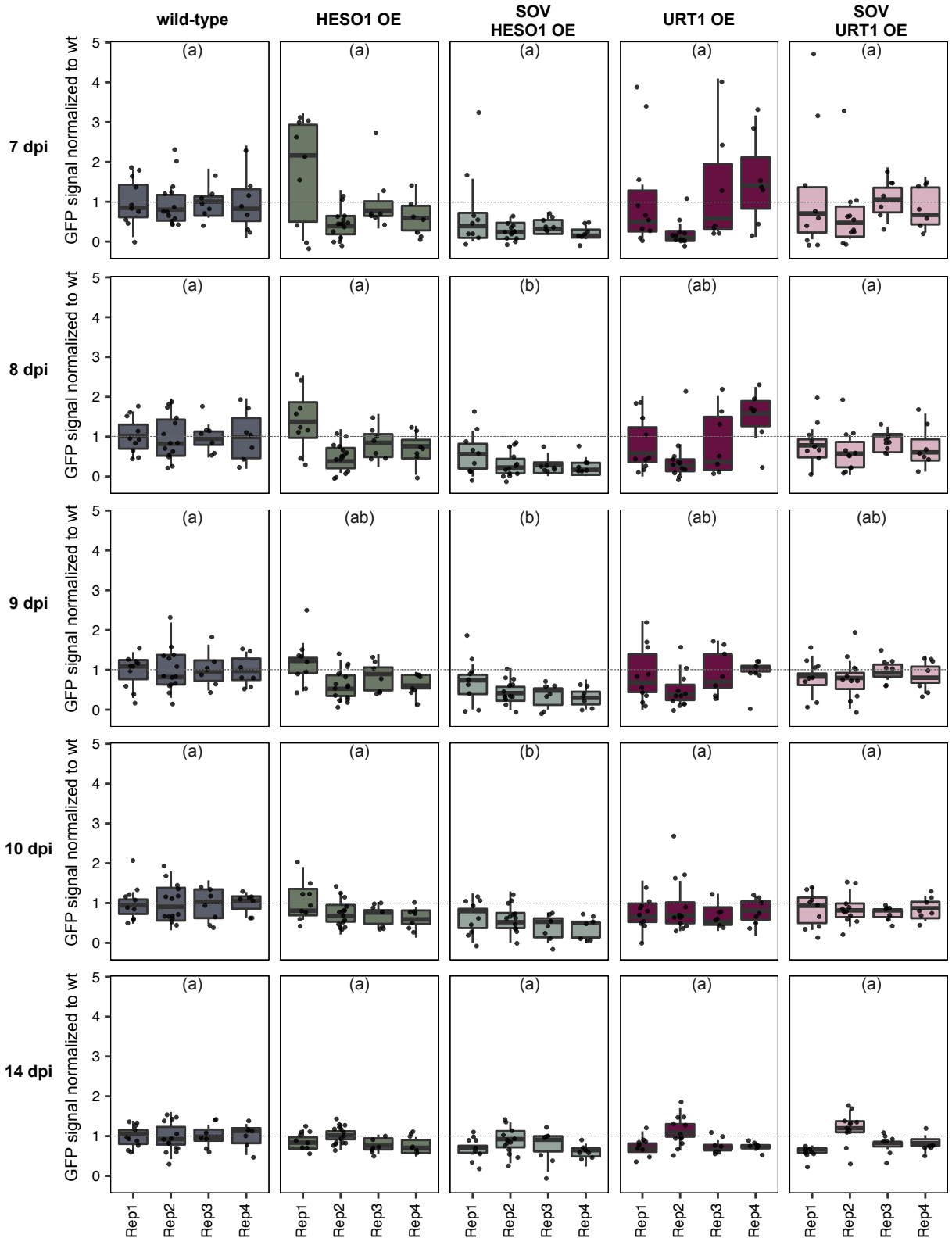


Figure 25: HESO1 overexpression reduces the accumulation of TuMV. Boxplots representing the GFP signal normalized to the mean signal of the wild-type plants by replicate at 7, 8, 9, 10 and 14 dpi. 7-15 infected plants per genotype for 4 independent biological replicates have been quantified. HESO1 OE and URT1 OE indicate the overexpressing lines in the *heso1-4* and *urt1-1* backgrounds, respectively. SOV HESO1 OE and SOV URT1 OE are the overexpressing lines in the Col-0 SOV(R705P) background. Letters show significant differences (p-value <0.05). The complete statistical analysis is shown in supplemental table 5. The grey line indicates the fold-change of 1, *i.e.*, no change compared to wild-type.

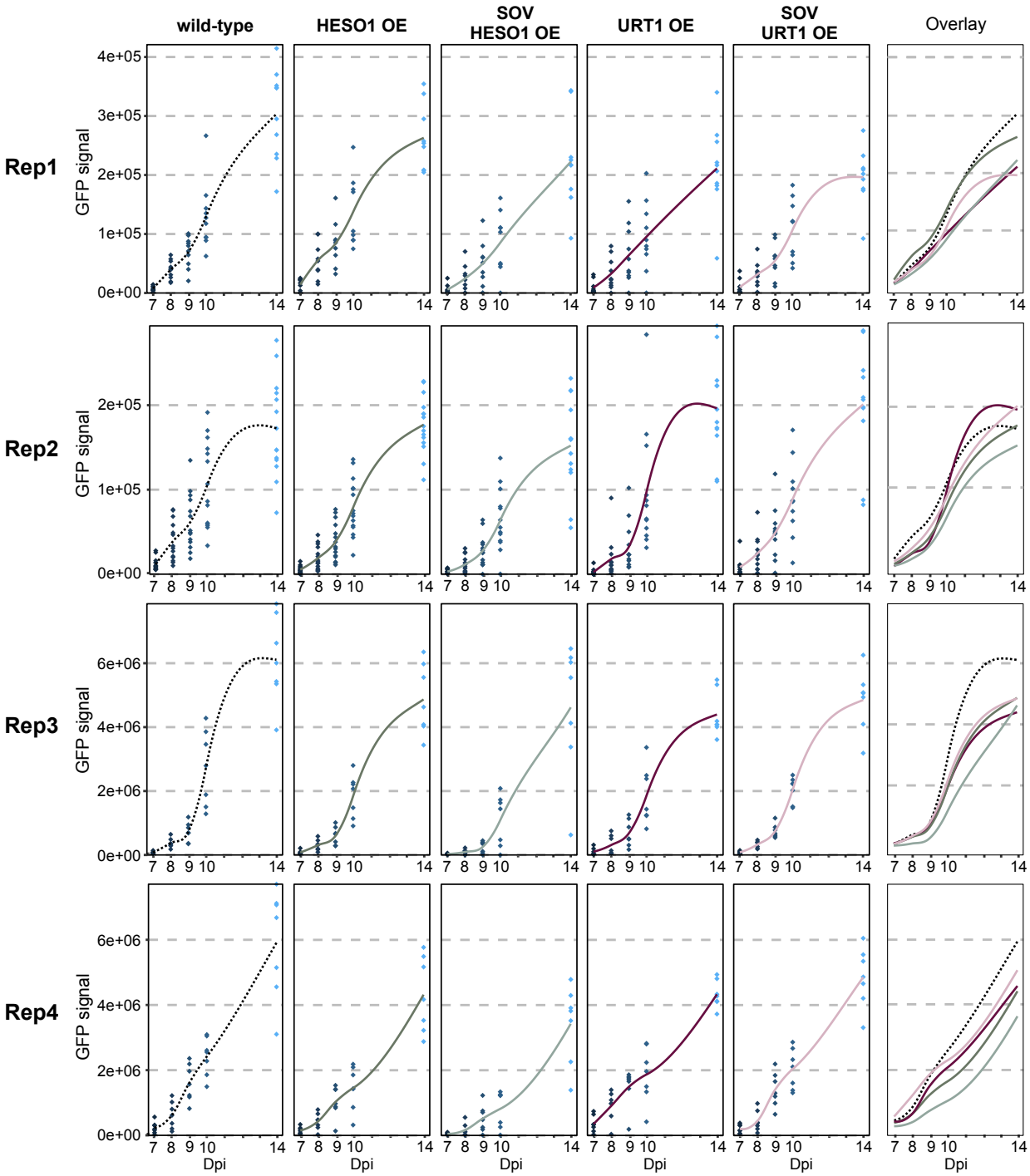


Figure 26: Slower TuMV infection upon TUTase overexpression. Plots representing the GFP signal measured in each plant at 7, 8, 9, 10 and 14 dpi (points) and the smoothed regression line (method = loess) by genotype and biological replicate. The last pannels show the smoothed lines of each genotype by replicate: wild-type (black dotted line), HESO1 (dark purple), SOV HESO1 (pink), URT1 (dark green), SOV URT1 (light green). 7-15 infected plants per genotype for four independent biological replicates have been quantified. HESO1 OE and URT1 OE indicate the overexpressing lines in the *heso1-4* and *urt1-1* backgrounds, respectively. SOV HESO1 OE and SOV URT1 OE are the overexpressing lines in the Col-0 SOV(R705P) background.

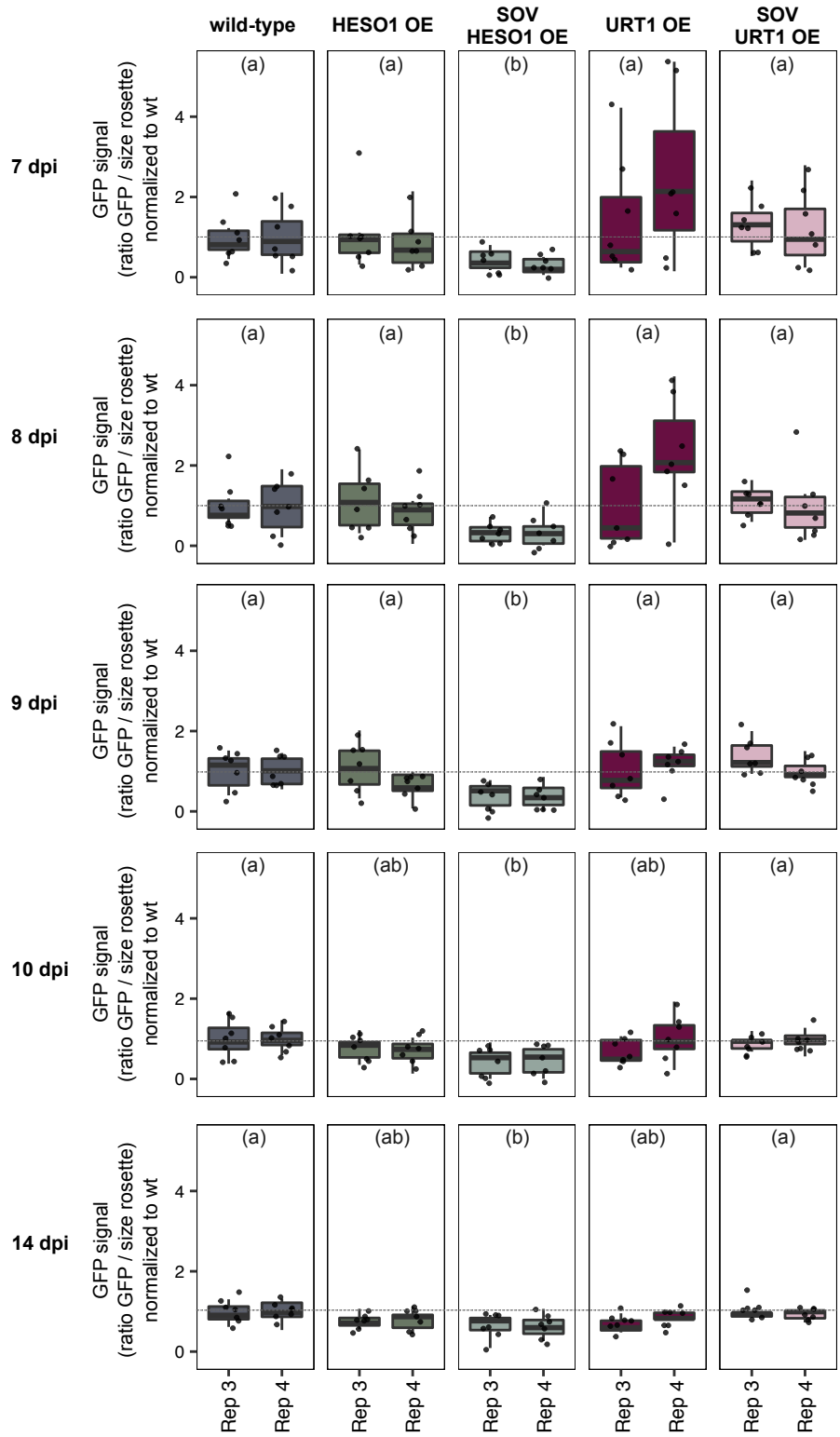


Figure 27: HESO1 overexpression reduces the accumulation of TuMV (ratio GFP signal / plant size). Boxplots representing the ratio (GFP signal / size of the rosette) for each plant, normalized to the mean ratio of the wild-type plants by replicate at 7, 8, 9, 10 and 14 dpi. 7 infected plants per genotype for four independent biological replicates have been quantified. HESO1 OE and URT1 OE indicate the overexpressing lines in the *heso1-4* and *urt1-1* backgrounds, respectively. SOV HESO1 OE and SOV URT1 OE are the overexpressing lines in the Col-0 SOV(R705P) background. Letters show significant differences (p-value <0.05). The complete statistical analysis is shown in supplemental table 5. The grey line indicates the fold-change of 1, *i.e.*, no change compared to wild-type.

Of note, while our team previously characterized the URT1 interactome in Arabidopsis flowers (Scheer et al., 2021), the HESO1 interactome has not been investigated in *planta* so far. The HESO1-tagged lines that I have produced during my PhD work will be further used to elucidate and compare TUTase interactomes, notably in the context of viral infection. Immunoprecipitation experiments followed by mass spectrometry will be performed using the different lines expressing the GFP- or the 4xmyc-tagged version of HESO1 in *heso1-4*. The analysis of transgenic lines with two different tags will allow to strengthen results to reduce background and non-specific interactants.

b. TuMV accumulation is reduced upon HESO1 overexpression

To analyze the viral accumulation in wild-type and TUTase overexpressing plants, I quantified the GFP signal in plants infected with TuMV-GFP for seven to fifteen plants per genotype in four biological replicates. For each replicate, I analyzed the fluorescence in a batch of plants from the start of the systemic infection at 7 dpi until 14 dpi. Even though there is variability in the signal intensities between the different plants of each genotype (Fig. 25 and Fig. S3-6), our data point out differences between the overexpressing lines as compared to wild-type.

First, at an early stage of infection (7-8 dpi), whereas all the wild-types of the four replicates are infected, as indicated by the presence of a GFP signal, several plants of the TUTase overexpressing lines do not have a green fluorescent signal (Fig. S3-6). This effect is more pronounced in the Col-0 SOV (R705P) HESO1-GFP plants with 20 to 50% of the plants that do not display a GFP signal.

Second, the signal is globally less intense when HESO1 is overexpressed in the Col-0 SOV(R705P) background (Fig. 25): from 8 to 10 dpi, the GFP signal is significantly lower in these plants, as compared to wild-type. In *heso1-4* background, we observed the same downward trend for replicates 2 to 4. Yet, the decrease is less pronounced and the replicate 1 behaves differently. More replicates will have to be analyzed to check the potential decrease of viral RNA accumulation in HESO1 overexpressing lines. Noteworthy, the decrease of GFP signal intensity is maximum before 10 dpi.

By contrast, no significant variation of GFP intensity was detected for the plants overexpressing URT1. Between 7 and 9 dpi, there is more biological variability between all plants and independent replicates than for HESO1 (Fig. 25 and Fig. S3-6), which could explain the difficulties to reveal a significant effect. Altogether, our data are not robust enough to conclude on an effect of URT1 overexpression in the accumulation of the GFP and thus, viral accumulation. The only robust and significant effect on the global TuMV-GFP accumulation is observed in plants expressing HESO1 in Col-0 SOV (R705P) background.

In addition, our time course analysis provides information on the evolution of the infection in *planta*, since the quantification is performed on the same plant from 7 to 14 dpi. To assess the speed at which the GFP signal increases over time in the different genotypes, I plotted the GFP signal in each plant over the measured time points for each genotype and biological replicate. In three out of four replicates, the GFP signal increases faster in wild-type than in the mutants (Fig. 26, grey curve). In addition, we see that the GFP signal evolution is slower, when compared to wild-type, in the four replicates for the plants overexpressing HESO1 in the Col-0 SOV(R705P) background (Fig. 26, dotted line). The effects in the other mutants *i.e.*, HESO1 overexpressed in *heso1-4* and URT1 in both genetic backgrounds, are

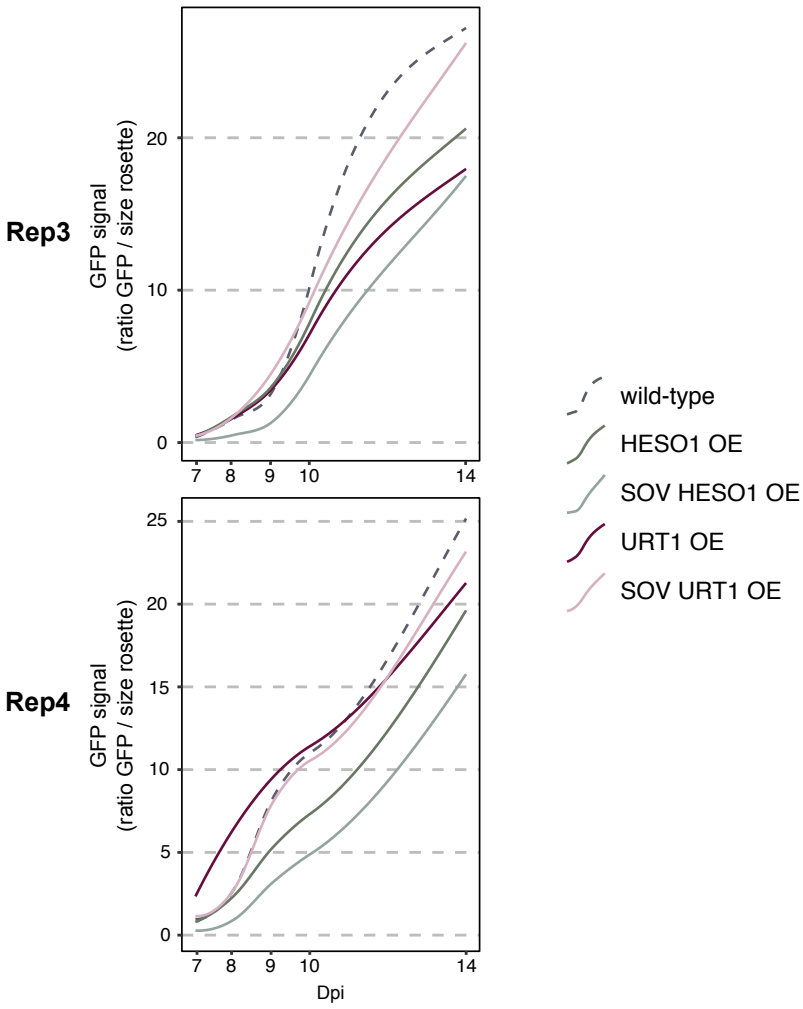


Figure 28: Slower TuMV infection upon TUTase overexpression (ratio GFP signal / plant size). Plots representing the smoothed regression line (method = loess) by genotype and biological replicate for the ratio (GFP signal / size of the rosette) for each plant at 7, 8, 9, 10 and 14 dpi: wild-type (grey), HESO1 OE (dark green), SOV HESO1 OE (green), URT1 OE (dark pink), SOV URT1 OE (light pink). 7 infected plants per genotype for 2 independent biological replicates have been quantified. HESO1 OE and URT1 OE indicate the overexpressing lines in the *heso1-4* and *urt1-1* backgrounds, respectively. SOV HESO1 OE and SOV URT1 OE are the overexpressing lines in the Col-0 SOV(R705P) background.

not as robust and replicable from one replicate to another. These data are in line with previous results (Fig. 25) and indicates that upon overexpression of HESO1 in the Col-0 SOV(R705P) background, the speed accumulation of the GFP signal is decreased compared to wild-type, suggesting a slower viral replication. This hypothesis is supported by the observation that the appearance of the GFP signal is delayed for plants of the Col-0 SOV (R705P) HESO1 line infected with TuMV.

Assessing the viral accumulation by quantifying the GFP has shown high variability in the level of infection between the plants of a same replicate. The analysis of the four replicates is performed by quantifying the pixel intensity of the green signal. This strategy does not take into account the size of the plants. Yet, young smaller plants usually are more sensitive to the virus, and a variation in plant sizes can bias the comparison between plants and/or between two genotypes. To try to improve the method of quantification and take the size of the plant into account, I also quantified the size of the whole plant for the biological replicates 3 and 4. Thus, I obtained two types of measurements: the raw intensity of the signal, that I previously used in Figure 25 and 26 and compared with the replicates 1 and 2, and the size-normalized measurements, for which I calculated the proportion of the rosette that emits green fluorescent signal. This strategy has the advantage to provide percentage values, that are more intuitive to compare between different biological replicates. I plotted the proportion of the GFP signal normalized to the mean of the wild-type for each plant by genotype and replicate (Fig. 27). We notice that we obtain the same tendencies regarding the GFP signal accumulation. In the plants overexpressing HESO1 in the Col-0 SOV (R705P) background, the ratio GFP signal / size of the rosette is significantly lower than in wild-type from 7 to 14 dpi. The results for the plants overexpressing URT1 in *urt1-1* or Col-0 SOV (R705P) backgrounds remain disparate, especially at 7 and 8 dpi, so, it is difficult to draw a conclusion on the impact of URT1 on the GFP signal accumulation.

The evolution of the ratio GFP / size of the rosette displays a similar profile as with the other quantification method (Fig. 28). It increases slower in Col-0 SOV (R705P) HESO1-GFP than in the other mutants and the wild-type. Even though there is still biological variability to consider, the quantification of the GFP signal related to rosette size allows to get rid of the potential variability due to plant growth and to limit bias.

Our analysis of TuMV accumulation has revealed that HESO1 overexpression results in slower viral infection. Our data on URT1 are less robust and display more variability. Altogether, these results are promising and could indicate the role of HESO1 in downregulating viral accumulation.

c. Preliminary data on TCV reduced accumulation upon TUTase overexpression

In Chapter 1.2 of the Results, we show that TCV RNA is uridylated at 10% (Article Fig. 4A), which is quite high for an RNA in plants. Indeed only 20% of Arabidopsis mRNAs are uridylated at level higher than 10% (Scheer et al., 2021). Both URT1 and HESO1 are involved in this uridylation activity and we detected a high proportion of truncated TCV RNAs that are uridylated (Article Fig. 4).

To determine whether TUTase proteins could have an impact on TCV viral accumulation, I performed preliminary experiments to analyze TCV infection in the TUTases overexpressing lines. For two independent replicates, I followed the emergence of systemic symptoms from 6 to 13 dpi and I harvested material at 13 dpi for six plants per genotype. TCV systemic infection in Arabidopsis is

Figure 29: Symptoms of systematically TCV-infected wild-type plant. Pictures of the mock and one symptomatic plant infected with TCV at 13 dpi for wild-type plants.

characterized by crinkled leaves at the center of the rosette and young leaves chlorosis (Fig. 29). At 6 dpi, no plants present any of these symptoms. At 7 dpi, at least one plant per genotype is systematically infected for both replicates (Table 5). Interestingly, all wild-type plants show infection symptoms at 9 dpi in replicate 1 and at 7 dpi in replicate 2 (Table 5). By contrast, it takes longer for the whole series of plants overexpressing one of the TUTase to be systematically infected, especially in the case of HESO1 in replicate 1 (Table 5): the time needed for the symptom appearance for the whole series was 13 dpi and 8 dpi in replicate 1 and replicate 2, respectively. This preliminary result suggests that TCV systemic infection could be delayed upon overexpression of HESO1. Further infections and viral RNA quantifications are required to consolidate this observation and to understand whether Arabidopsis TUTase expression could impact TCV RNA destabilization.

Biological replicate #1

Genotype	6dpi	7dpi	8dpi	9dpi	10dpi	13dpi
Wild-type	no symptoms	5/6	5/6	6/6	6/6	6/6
urt1-1 URT1-GFP		3/6	4/6	5/6	5/6	6/6
heso1-4 HESO1-GFP		2/6	4/6	5/6	5/6	5/6
Col-0 SOV(R705P) URT1-GFP		5/6	5/6	6/6	6/6	6/6
Col-0 SOV(R705P) HESO1-GFP		2/6	4/6	4/6	4/6	5/6

Biological replicate #2

Genotype	6dpi	7dpi	8dpi	9dpi	10dpi	13dpi
Wild-type	no symptoms	6/6	6/6	6/6	6/6	6/6
urt1-1 URT1-GFP		6/6	6/6	6/6	6/6	6/6
heso1-4 HESO1-GFP		4/6	6/6	6/6	6/6	6/6
Col-0 SOV(R705P) URT1-GFP		5/6	6/6	6/6	6/6	6/6
Col-0 SOV(R705P) HESO1-GFP		3/6	6/6	6/6	6/6	6/6

Table 5: Monitoring of systemic symptoms in TCV-infected plants. The number of plants presenting systemic symptoms are indicated over the total number of plants analyzed for two independent replicates from 6 to 13 dpi. Red writing highlights the timepoints at which not all the plants from a same batch have symptoms.

3. First insights on the relationship between uridylation and degradation of viral RNA in the cytosol

1. Selection of mutants impaired for RNA degradation

The key function of RNA uridylation, conserved in eukaryotes, is to promote RNA decay. Interestingly, uridylation has also been proposed to induce viral RNA decay in *C. elegans* and in mammalian cells (Le Pen et al., 2018). As a first experiment to investigate mechanisms underlying the degradation of uridylated viral RNA in plants, we tested the impact of inactivating proteins involved in 5'-3' and 3'-5' cytosolic RNA degradation on TuMV accumulation and uridylation. The first set of mutants

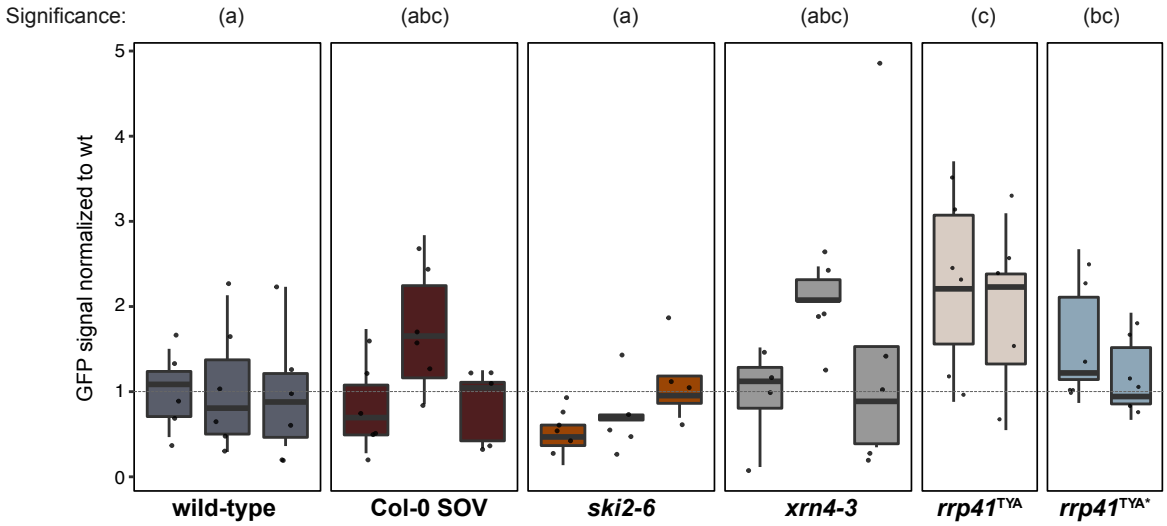


Figure 30: Erratic effect of mutating cytoplasmic decay factors on TuMV accumulation. Boxplots representing the GFP signal normalized to the mean signal of the wild-type plants by replicate at 10 dpi. 4-6 infected plants per genotype for 2 or 3 independent biological replicates have been quantified. Col-0 SOV corresponds to the Col-0 SOV(R705P) transgenic line. Each boxplot represents one biological replicate. Letters show significant differences (p-value <0.05). The complete statistical analysis is shown in supplemental table 5. The grey line indicates the fold-change of 1, *i.e.*, no change compared to wild-type.

were selected to address a potential role for the 5'-3' exoribonuclease XRN4, the 3'-5' exoribonuclease SOV, the exosome cofactors SKI2 and the phosphorolytic activity of the core exosome subunit RRP41.

Several studies explored the impact of the 5'-3' exoribonuclease XRN4 on TuMV RNA accumulation but their results are rather contradictory. One study reported a mutation in XRN4 that results in a decreased susceptibility to TuMV (Vogel et al., 2011), while others show that XRN4 negatively impact TuMV accumulation (Li and Wang, 2018; Ma et al., 2019). Hence, our analysis is relevant to better understand how XRN4 impacts TuMV accumulation.

We also wanted to explore the potential role of the 3'-5' decay pathways in viral RNA degradation. Uridylation has been reported to stimulate 3'-5' decay by DIS3L2 in mammals (Łabno et al., 2016; Lubas et al., 2013; Reimão-Pinto et al., 2016; Thomas et al., 2015; Ustianenko et al., 2013, 2016). So, to determine whether SOV, the DIS3L2 homolog in plants, could be involved in viral RNA decay and whether uridylation is an enhancer of its activity, we compared wild-type Col-0 plants (in which SOV is inactive, or at least severely impaired) and Col-0 plants expressing SOV (R705P).

We also used the null allele *ski2-6*, in which the cytosolic exosome cofactor SKI2 is mutated. SKI2 is a cytoplasmic RNA helicase, which is part of the SKI complex *via* its interaction with the tetratricopeptide repeat protein SKI3 and a dimer of the WD40-repeat protein SKI8/VIP3 (Lange and Gagliardi, 2022). SKI2 assists the exosome in cytoplasmic decay in mammals and plants (Branscheid et al., 2015; Lange and Gagliardi, 2022; Tuck et al., 2020). To date, there is no data on the involvement of the SKI complex and the exosome in viral RNA decay.

Finally, we analyzed a mutant impaired for a subunit of Exo9, RRP41. Arabidopsis Exo9 possesses a distributive phosphorolytic activity, harbored by RRP41 (Sikorska et al., 2017). Our team has demonstrated that mutating the phosphate coordination (*rrp41*^{TYA} mutation) and both phosphate and magnesium coordination (*rrp41*^{TYA*} mutation) within RRP41 abolishes Exo9 phosphorolytic activity (Sikorska et al., 2017). Arabidopsis Exo9 can also catalyze the reverse reaction *i.e.*, it is able to synthetize RNA tails using nucleoside diphosphates (Sikorska et al., 2017), like its archaeal counterpart and bacterial PNPases (Lange and Gagliardi, 2022). Heteropolymeric tails, that looks like those detected on exosome RNA substrates, have been detected on several viral RNAs (Huo et al., 2016; Li et al., 2014). Therefore, we wanted to assess whether such heteropolymeric tails could be found on TuMV RNA and whether the exosome could be involved in their synthesis.

2. Impact of mutating decay factors on TuMV accumulation

To investigate the impact of cytoplasmic RNA decay factors on TuMV infection, we analyzed the accumulation of the GFP signal at 10 dpi in wild-type, Col-0 SOV (R705P), *xrn4-3*, *ski2-6* for four to six plants per genotype for three independent biological replicates and in *rrp41*^{TYA} and *rrp41*^{TYA*} for five to six plants per genotype for two replicates. The GFP signal intensities have been normalized to the mean of the wild-type values. There is no difference in the GFP signal between wild-type and Col-0 SOV (R705P), *ski2-6* and *xrn4-3* for the three replicates (Fig. 30). We can observe a significant increase in the signal intensity for *rrp41*^{TYA} and *rrp41*^{TYA*} mutants, as compared to wild-type (Fig. 30). This observation is interesting but requires further work to be consolidated. So far, I cannot explain the difference between the two *rrp41* mutants. Considering the biological variability among the infected

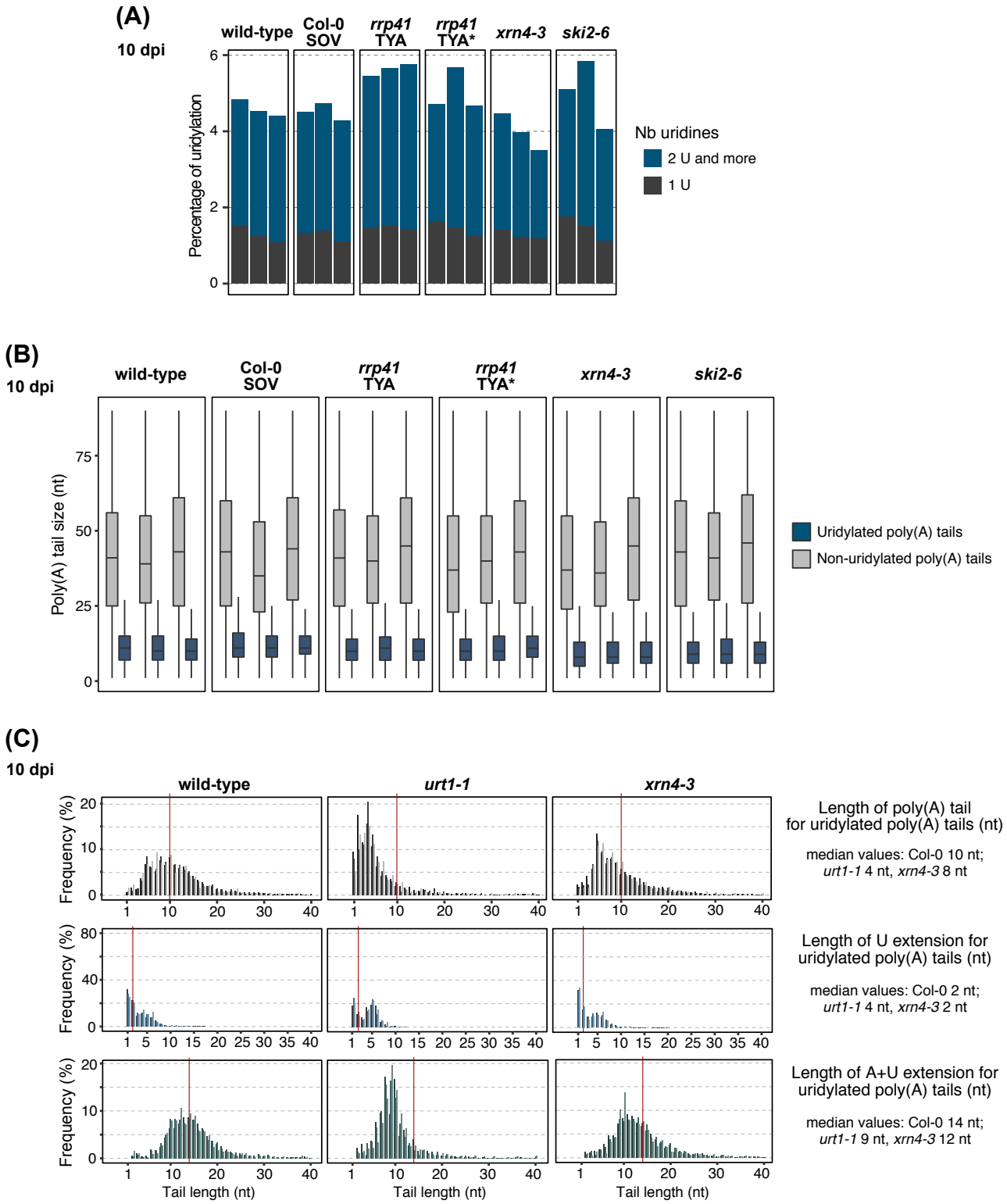


Figure 31: Uridylation takes place on deadenylated TuMV RNA. (A) Barplot showing the percentage of uridylated reads (polyAandU and polyU classes), calculated with the total number of reads for TuMV RNA in each plant as denominator. The color code indicate the number of uridines detected (1 U dark grey and 2 Us and more blue). **(B)** Boxplots showing the poly(A) tail size of uridylated tails (blue) and non-uridylated tails (light grey) for each biological replicate. **(C)** Distribution of the extensions for uridylated poly(A) tails shown for 1 to 40 As: poly(A) tail size of uridylated reads (black/grey), length of the uridine extensions (blue), total length of the extension (poly(A) + U-tail, green). The distribution has been calculated using as denominator the number of uridylated polyadenylated reads for TuMV in each replicate for poly(A) tails ranging from 1 to 89 As. The red line indicates the median in wild-type. On each barplot, each bar represent one biological replicate.

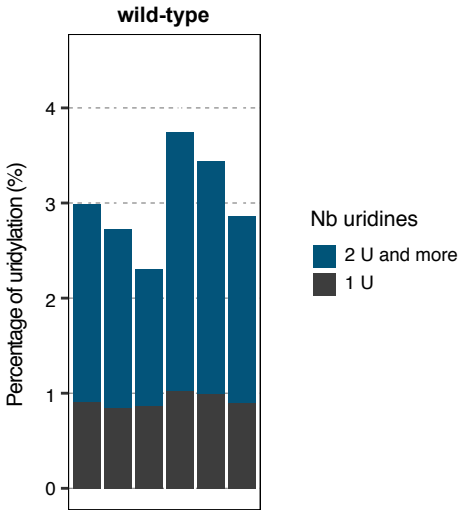


Figure 32: Percentage of uridylated polyadenylated TuMV RNA in wild-type at 14 dpi. Barplot showing the percentage of uridylated polyadenylated reads, calculated using the total number of reads for TuMV RNA in each plant as denominator. The color code indicate the number of uridines detected (1 U, dark grey and 2 Us and more, blue). Each bar represent one infected plant. Three plants of two independent replicates have been analyzed.

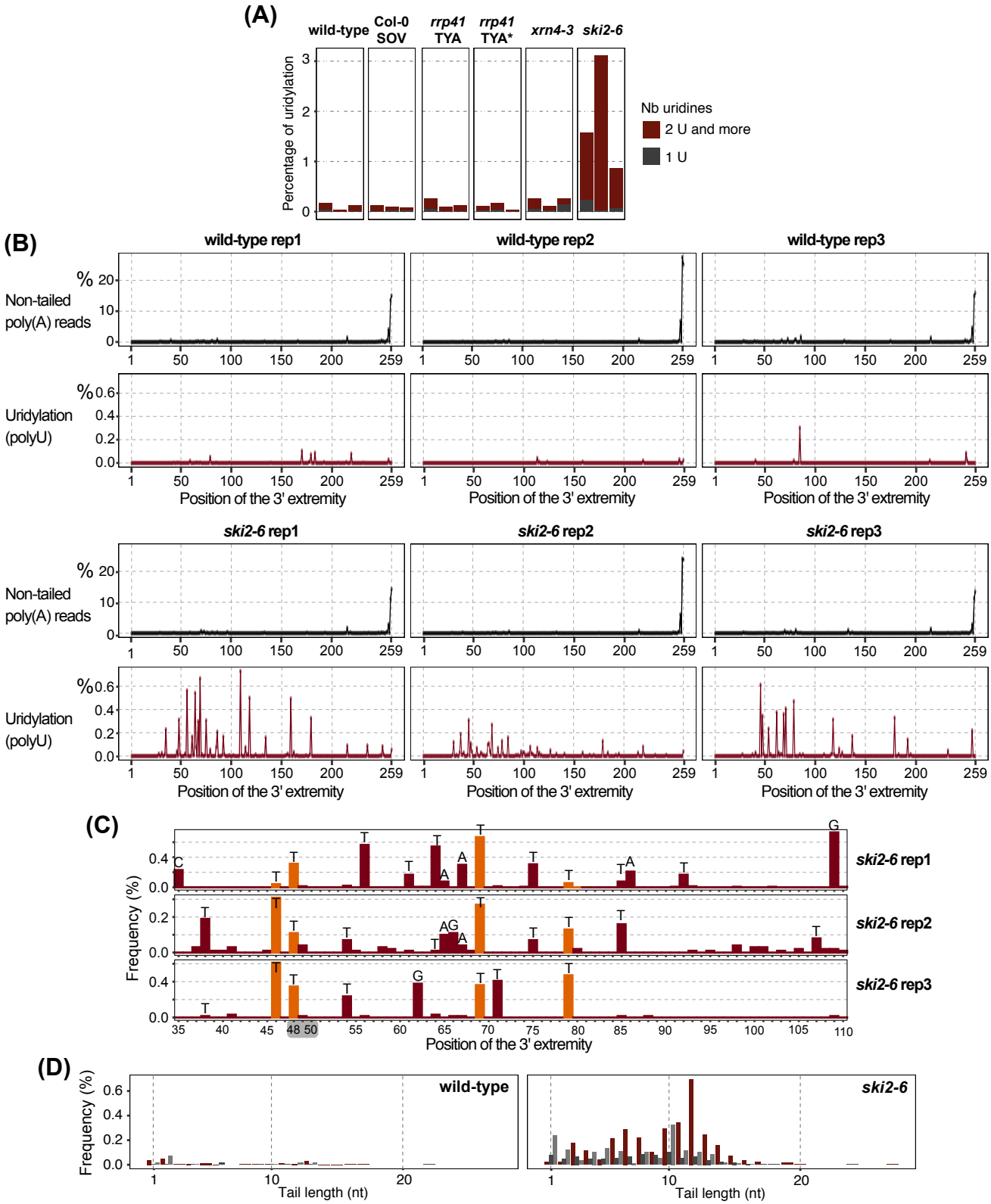


Figure 33: Accumulation of truncated uridylated TuMV RNA in the absence of SKI2. (A) Barplot showing the percentage of uridylated non-polyadenylated reads. The color code indicate the number of uridines detected (1 U dark grey and 2 Us and more red). **(B)** High resolution mapping of TuMV RNA 3' extremities in wild-type and *ski2-6* plants. Frequencies of reads at each 3' end position are shown for polyadenylated reads (black) and U-tailed non-polyadenylated reads (dark red). The positions 1 and 259 correspond to the first and the last nucleotide of the PCR amplicon, respectively, the last nucleotide being considered as the mature extremity of the RNA. **(C)** Close-up view of the high resolution mapping of TuMV 3' extremities *ski2-6* plants. Frequencies of reads mapping from positions 35-100 are shown for U-tailed reads. The position on the reference sequence is indicated below the graph and the corresponding nucleotide is shown on top of some bars. The UAG stop codon is indicated by a light grey square. The conserved uridylated positions between the three replicates are colored in orange. **(D)** Barplot showing the length distribution of the uridine extensions in wild-type and *ski2-6*. On each barplot, each bar represent one biological replicate. The percentage in (A) and the frequencies in (B), (C) and (D) have been calculated with the total number of TuMV RNA reads in each plants as denominator.

plants (Fig. S7), this result needs to be confirmed by the analysis of more plants and further biological replicates.

3. Uridylation is detected on deadenylated TuMV RNAs

To further investigate the potential function of uridylation in the degradation of TuMV RNA, we monitored the 3' modifications on TuMV RNA by 3' RACE-seq in three biological replicates of wild-type and degradation/uridylation mutants at 10 dpi. One replicate consists of 2-3 plants pooled together. TuMV RNA being polyadenylated, we can distinguish in our 3'RACE-seq data, two types of uridylation: uridine addition at the poly(A) tail and uridine addition to non-polyadenylated 3' extremities, corresponding to truncated viral RNAs. We observed that the global percentage of uridylation (U-tail + uridylated poly(A) tails) in wild-type at 10 dpi is similar to that at 14 dpi (Article Fig. 6A. Fig 20). The percentage of uridylation on polyadenylated TuMV RNA in wild-type is also about 4% at 10 dpi, like it is at 14 dpi (Fig. 31A and 32). We can also observe that there is no difference in the uridylation percentage of TuMV polyadenylated RNAs in the mutants for cytosolic decay factors (Fig. 31A).

As observed in wild-type and TUTase mutants at 14 dpi (Article Fig. 6B, C), uridylation of TuMV RNA is detected on short poly(A) tails in wild-type and cytosolic decay mutants at 10 dpi (Fig. 31B). Interestingly, the distribution of the poly(A) tails uridylated in *urt1-1* and *xrn4-3* is shifted toward shorter tails, especially in *urt1-1* (Fig. 31C). Indeed, the median length of the poly(A) tail for uridylated poly(A) tails is 10 As in wild-type plants whereas it drops to 4-5 As and 8 As in *urt1-1* and *xrn4-3*, respectively (Fig. 31C). In line with this observation, the total length of the uridylated poly(A) tail (A+U extension) is shorter in *urt1-1* and *xrn4-3* compared to wild-type (Fig. 31C). In the absence of URT1, the remaining uridylation, probably catalyzed by HESO1, is detected on shorter poly(A) tails than in wild-type. So, this further demonstrates the distinct substrate specificities between URT1 and HESO1 for viral RNAs.

Most U-tails detected on TuMV poly(A) tail in wild-type range from 1 to 6 nucleotides (Fig. 31C). As previously mentioned in Chapter 1, there is an increased proportion of U-tails between 4 and 6 Us in *urt1-1*, as compared to wild-type (Fig. 21B, C and 31C). This means that in the absence of URT1, HESO1 adds longer uridine extensions to TuMV RNA poly(A) tail. Of note, uridine extensions detected on TuMV RNA in *xrn4-3* are similar to wild-type (Fig. 31C). This means that even if the proportion of deadenylated TuMV RNA increases, the length of the uridine extension does not change, which is different from the behavior observed on mRNAs by our team where uridylation repairs deadenylated mRNAs (Zuber et al., 2016).

Our results on TuMV RNA uridylation in cytosolic decay mutants confirmed the similarities with mRNA uridylation to some extents, but it also revealed differences regarding the patterns of uridylation in the absence of XRN4.

4. Accumulation of uridylated truncated viral RNAs in ski2-6

In the previous paragraph, we discussed uridylation on polyadenylated TuMV RNA. Our data has also revealed an accumulation of uridylated truncated TuMV RNAs in *ski2-6* as compared to wild type (Fig. 33A). Indeed, whereas most non-tailed and tailed reads map to the mature 3' extremity of the RNA,

a high proportion of the reads that are modified with non-templated uridines map within the reference sequence, meaning that they correspond to truncated TuMV RNAs (Fig. 33B).

In addition, we observed a cluster downstream of the stop codon of TuMV polyprotein, between the positions 35 and 110 which is enriched in reads that are modified with non-templated uridines (Fig. 33C). We observed a high variability in the positions that are uridylated between the three replicates. Nevertheless, there are four positions that we retrieve in the three replicates. Two of them are close to the stop codon and the others are located within the 3' UTR (Fig. 33C, orange bars). This could indicate that viral RNA fragments with these extremities are generated by 3'-5' trimming followed by uridylation and that they are stabilized in the absence of SKI2. Moreover, most of these positions end with a T in the sequencing data, which corresponds to a uridine in TuMV RNA sequence (Fig. 33C).

Interestingly, these uridine extensions are much longer than the extensions we detected after the poly(A) tail, with most of them ranging from 5 to 14 Us (Fig. 33D). This result suggests that these long uridine extensions are preferentially added to a 3' terminal uridine. In light of the results previously presented, HESO1 is a good candidate to add these long uridine extensions.

We show that uridylated truncated TuMV RNAs accumulate in the absence of the exosome cofactor SKI2. This suggests that under wild-type conditions, these fragments are so little abundant and unstable that we cannot detect them, indicating that they are undergoing decay. This result paves the way to the better understanding of viral RNA decay via the cytosolic exosome.

Discussion and perspectives

1. Limitations of 3'RACE-seq and alternative approaches

The first objective of my thesis project was to explore the diversity of the uridylation landscape for plant viral RNAs. I analyzed 3' terminal uridylation by 3'RACE-seq, an in-depth strategy that combines cDNA nested PCR amplification with Illumina-type high-throughput sequencing. In this part, I will discuss the limitations of this strategy regarding our project and propose alternative approaches.

1. Limitations of 3'RACE-seq

3'RACE-seq relies on the ligation of an adapter to the RNA 3' extremity, followed by nested cDNA PCR amplification. The adapter is adenylated at its 5' end, providing the adenyl group required by the T4 RNA ligase so, no ATP is needed in the ligation reaction. This adapter (Chang et al., 2014) is composed of a 5 nt-long delimiter sequence, a 15 nt-long random region and a region onto which the RT primer anneals to initiate cDNA synthesis. The delimiter sequence allows to distinguish the adapter sequence from the RNA tail. The random region is used during the analysis in a deduplication step: identical random regions are removed so, we analyzed individual RNA molecules and not PCR duplicates. The 3' end of the adapter is blocked by a dideoxynucleotide to prevent self-ligation. Nested-PCR allows to specifically amplify the target of interest and the addition of barcodes to multiplex samples. The libraries are paired-end sequenced by Illumina chemistry, which provides good cover and depth.

The pipeline of analysis created by Hélène Zuber offers three options to analyze 3' tailing on targets RNAs (see Material and Methods). For non-polyadenylated RNAs, the sequence of reference is used to map the read 2, so we can map at a high resolution the 3' extremities of the RNA and potential trimmed molecules and determine non-templated nucleotides at the 3' extremities. For polyadenylated RNAs, either the read 2 is mapped onto the reference sequence and 3' non-templated nucleotides are determined, or a poly(A) tail longer than 8 As is searched and 3' modifications on the poly(A) tail are assessed. This pipeline allows the broad study of 3' tailing on target RNAs.

In the context of this study, there are however some elements regarding 3'RACE-seq that have to be considered. First, the ligation step requires the RNA to end with a 3' hydroxyl group, accessible to the ligase. Hence, modified nucleotides or particular structures could prevent the binding and subsequent ligation reaction by the ligase. This could particularly be the case for viral RNAs that end with a TLS or other strong stable structures.

In the particular case of polyadenylated viral RNAs, 3'RACE-seq displays several biases regarding poly(A) tail length measurement. First, poly(A) tails size estimation with Illumina chemistry can be challenging because of the difficulty to read homopolymeric sequences. Indeed, (Chang et al., 2014) reported an overestimation of poly(A) tail sizes when measuring long homopolymeric stretches. To counteract this biais, Chang and colleagues developed a complementary algorithm called Tailseeker

that uses fluorescence signals instead of the standard base calling software of Illumina to estimate the poly(A) tail size. Second, there is also a bias of the 3' RACE-seq strategy toward short poly(A) tails. This biais is linked to the PCR amplification step, as short poly(A) tails are more efficiently amplified than long, but is also largely caused by the Illumina sequencing. Indeed, as previously reported in (Morgan et al., 2017), we observed a lower clustering efficiency for long poly(A) tails when compared to short poly(A) tails. This can be revealed when analyzing the control spike-ins used on each sequencing run: the recovery of spike-ins with no or short poly(A) tail (Poly A Spike_8 and Poly A Spike_N, see Material and Methods) is systematically higher than for the spike-in with 80 As.

In addition, our strategy does not allow to distinguish between the different viral RNA species (genomic, subgenomic, defective interferent if it exists). So, we cannot affirm that our viral RNA tailing analysis concerns the genomic RNA, except for the experiment in which we analyzed encapsidated GFLV RNAs. Indeed, among the viral species that we analyzed, all but TuMV and species of the *Secoviridae* have been reported or suggested to produce sgRNAs. This has been suggested for grapevine-infecting viruses GRGV, GRSPaV, GPGV and GVB but there is no experimental data yet. As sgRNAs and gRNAs can have identical 3'UTRs, our 3'RACE-seq primers can anneal to both RNA species. Therefore, it is in fact highly probable that we mapped other viral species because sgRNAs accumulate at higher or similar levels than gRNA, as it has been illustrated by Northern blot analyses of TYMV, PVX, BaMV, *Pelargonium line pattern virus* (*Tombusviridae*), *Potato leafroll virus* (*Solemoviridae*), and TBSV RNAs (Cho and Dreher, 2006; Garcia et al., 2014; Hung et al., 2014; Pérez-Cañamás et al., 2021; Tacke et al., 1990; Wu and White, 2007).

2. Alternative methods to overcome 3'RACE-seq limitations

To discriminate between gRNA and other viral RNA species present within the plant cell, it could be useful to perform a size selection prior to generate the libraries. For example, to enrich in gRNA, a biotinylated oligo specific for gRNA sequence could be used and all other viral RNAs that do not anneal to the primer could be discarded. This would however require to use a high quantity of total RNA and to adapt the enrichment procedure for each viral RNA of interest, individually. Another alternative that could be used on a total RNA sample, without needing to optimize it for each target, could be to purify the total RNA on beads by size selection and discard the smaller RNA species, for example sgRNAs. The most efficient and certain way to discriminate between gRNA and sgRNAs would be to sequence the whole RNA molecule.

Two long-read sequencing techniques could be used: Single Molecule Real-Time (SMRT®) sequencing (PacBio technology) or nanopore-based strategies proposed by Oxford Nanopore Technologies (ONT). Importantly, these two long-read sequencing techniques would also overcome the biaises associated to the estimation of long poly(A) tails linked to Illumina sequencing.

SMRT® sequencing is a DNA-sequencing method that offers the possibility of long high fidelity reads sequencing (up to tens of kilobases) with high accuracy (Eid et al., 2009). PacBio technology has been used to study poly(A) tail length in mouse oocytes and rat liver with PAIso-seq (Liu et al., 2019) and in human cells and *C. elegans* with FLAM-seq (Legnini et al., 2019).

Nanopore-based sequencing strategies allow to sequence cDNA, as Illumina or SMRT, but also directly the RNA. Direct RNA Sequencing (DRS) has become widespread in the last few years for analyzing mRNAs and measuring their poly(A) tail size. This strategy offers the possibility of long-read full RNA molecule sequencing without RNA fragmentation or amplification. A sequencing primer that carries a motor protein is attached to the RNA. The motor protein drives the RNA through an ion pore at consistent speed, so the dwell time of the poly(A) tail in the pore correlates with its length. DRS allows to remove some biases related to homopolymeric sequencing, PCR amplification and read length limitations. It is a useful tool to measure poly(A) tail length, as it has been published for human cells, mouse, viral RNAs, *A. thaliana, S. cerevisiae* and *C. elegans* (Bilska et al., 2020; Depledge and Wilson, 2020; Gewartowska et al., 2021; Kim et al., 2020b; Krause et al., 2019; Maier et al., 2020; Parker et al., 2020; Roach et al., 2020; Scheer et al., 2021; Shin et al., 2022; Tudek et al., 2021; Workman et al., 2019; Zhang et al., 2020b).

Yet, most of the DRS protocols used to date do not allow to analyze non-polyadenylated RNAs: the standard nanopore adapter, that is annealed to the RNA 3' end, contains an anchored oligo (dT) and the mRNA enrichment often rely on oligo(dT) beads. Another limitation is the important amount of RNA required for the preparation of DRS libraries (> 150µg). To overcome these limitations, other nanopore-based strategies have been developed and they rely on cDNA sequencing. FLEP-seq2 has been recently published to estimate poly(A) tail size in *A. thaliana* from cDNA sequencing (Long et al., 2021). This method is based on the Nanopore PCR-cDNA sequencing kit which uses ligase-free attachment of rapid sequencing adapter. FLEP-seq2 has recently been used to establish a dataset of poly(A) tail lengths in different Arabidopsis tissues (Jia et al., 2022). FLEP-seq2 allows to capture different types of RNAs regardless of their tail composition. It can provide RNA abundance measurements, poly(A) tail size estimation, and tail composition.

As compared to SMRT® sequencing, the error rate of nanopore sequencing remains high. Despite this, nanopore sequencing was the first strategy to offer either a reliable direct RNA sequencing. It grants full-length RNA or cDNA sequencing and less bias in the homopolymer sequencing. In addition, in one run, depending on the pipeline of analysis used on the data, information about poly(A) tail size but also differential expression, splicing isoforms, nucleotide modifications (in DRS) and terminal modifications can be obtained. This promising technology is in constant evolution and improvement, regarding its chemistry and the library generation, but also the pipelines of analysis.

2. 3' end modifications of viral RNAs in plants

1. Viral RNAs with known 3' TLS

a. 3' adenylation

In our analysis of viral RNA uridylation by 3'RACE-seq, we worked with several RNAs that display a tRNA-like structure (TLS) at their 3' extremity: TMV, PCV, TYMV and CMV RNAs. These TLS can be aminoacylated with histidine, valine or tyrosine, respectively. Of note, aminoacylation will prevent their

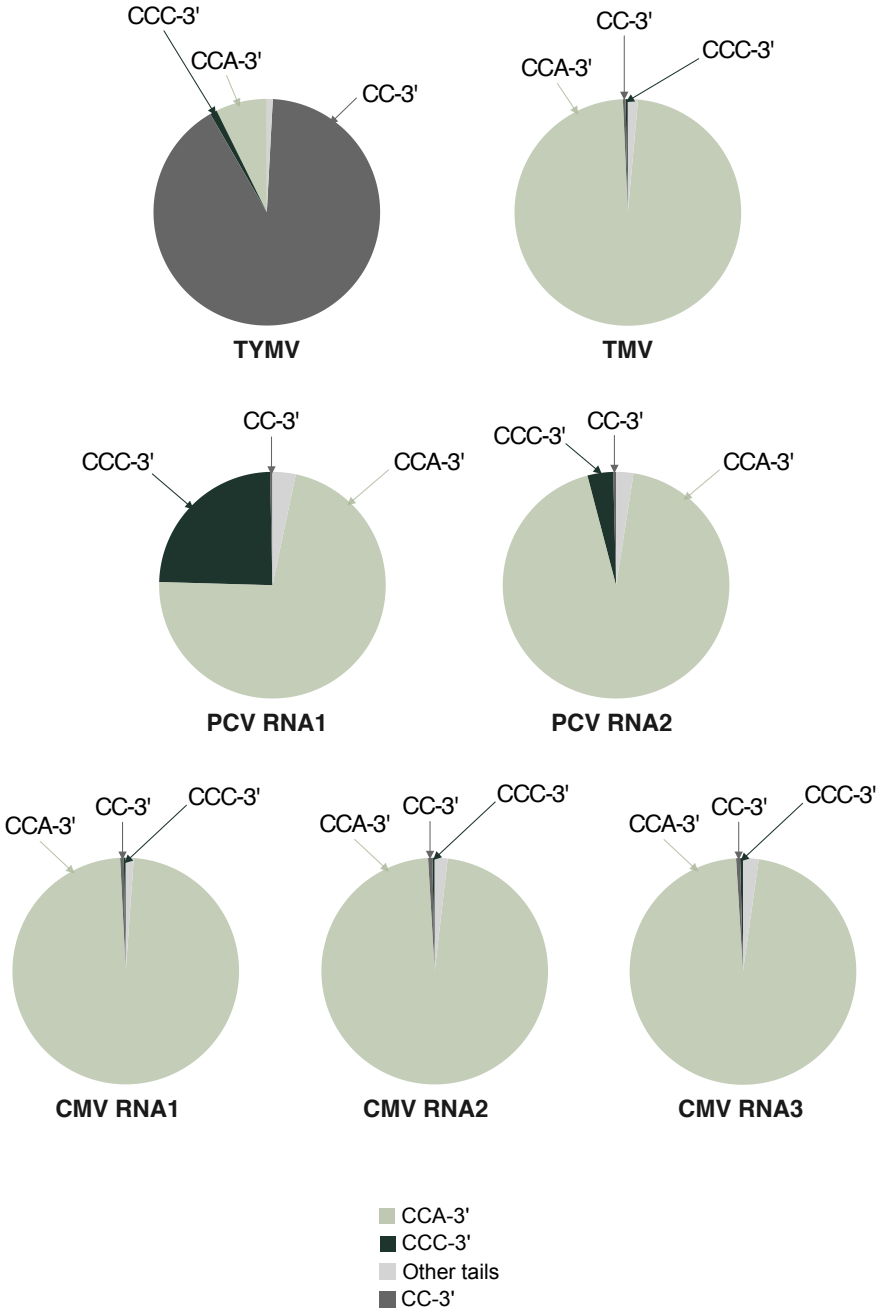


Figure 34: 3' extremities of plants viral RNAs ending with TLS determined by 3'RACE-seq. Piecharts representing the proportion of each type of 3' extremity for TLS-ending RNAs: CCA-3' (light grey), CC-3' (grey), CCC-3' (dark green), and other types of tails (CN-3' or CCN-3' with N corresponding to a U or a G, light grey). CCA-3' corresponds to the canonical ending of a TLS. The percentage of each 3' end has been calculated using as denominator the total number of reads for each viral RNA and each of the three independent plants.

detection by 3'RACE-seq so, the fraction viral RNAs ending with a TLS that we analyzed corresponds to a population that is not aminoacylated.

Like cellular tRNAs, TLSs can be adenylated by a CCA-adding enzyme (Dreher, 2010). Our results provide a report of the *in vivo* adenylation status for several plant viral TLS. We detected 3' adenylation mainly on TYMV TLS (Fig. 16A). In the population of non-aminoacylated TYMV RNAs detected by 3'RACE-seq, most of them end in CC_{OH}-3', with a non negligeable proportion that end in CCA_{OH}-3' (Fig. 34). By contrast, TMV, PCV and CMV RNAs end majorly in CCA_{OH}-3' (Fig. 34). Our results indicate that adenylation in *planta* can happen and that it can represent the majority of the RNA molecules that we analyzed for TMV, PCV, and CMV RNAs. However, it is not possible to say whether it is added randomly by the cellular CCA-adding enzyme when the TLS is accessible or whether it has become encoded by the viral genome.

b. 3' cytidylation

CCA_{OH}-3' is not the only sequence that we have detected at the 3' extremity of TLS-ending RNAs. Especially on PCV RNA1, mono-cytidylation is remarkably abundant and exclusively localized at the penultimate nucleotide of the TLS (Fig. 17). So, there is a high proportion of PCV RNA (25% RNA1 and 4% RNA2) that ends with CCC_{OH}-3' (Fig. 17E), by contrast with the other TLS-ending RNAs (Fig. 34). So far, we are not able to explain this singular terminal addition that we detected on PCV RNAs. *In vitro* tests with *E. coli* CCA-adding enzyme have shown that this enzyme is able to synthetize short poly(C) tails (Hou, 2000; Seth et al., 2002). In addition, CCA-adding enzymes have been unexpectedly reported to be capable of adding CCACCA extensions *in vitro*, for archaeal, *E. coli* and human enzymes, and also *in vivo* for *E. coli* and human CCA-adding enzymes (Betat et al., 2004; Wilusz et al., 2011). These data suggest that the activity of the CCA-adding enzyme is not limited to CCA addition to mature tRNAs. Hence, it is possible that the observed PCV RNA terminal cytidylation could be ensured by the cellular CCA-adding enzyme.

c. 3'uridylation

As compared to adenylation and cytidylation, terminal uridylation of TLS-ending viral RNAs has been detected at low level (Article Fig. 2). Our data indicate that viral RNAs ending with a TLS are not heavily modified with non-templated nucleotides, except for PCV RNA1 (Article Fig. S1; Fig. 17). Moreover, when we assessed the impact of the absence of URT1 and HESO1 on TYMV RNA uridylation level by 3'RACE-seq, we did not observe any substantial decrease: percentages with or without the TUTases are below 0.5% (Fig. 35). This suggests either that this low uridylation percentage corresponds to background noise or that URT1 and HESO1 are not involved in TYMV RNA uridylation. Nevertheless, it is likely that this RNA is difficult to access for the cellular TUTases because of the TLS, but also because of a probable competition with other 3' terminal modifying enzymes. Indeed, TYMV RNA has been described *in vitro* to be adenylated by the CCA-adding enzyme and aminoacylated by a cellular aminoacyl-tRNA synthetase. Although we do not know the proportion of TYMV TLS that are aminoacylated *in planta*, it is likely that these cellular factors are also able to modify TYMV RNA *in vivo*. As TLSs mimic cellular tRNAs, it is likely that the tRNA-modifying enzyme display a stronger affinity than TNTases. Hence, it is not surprising that viral RNAs with a 3'TLS are not uridylated at high level. There

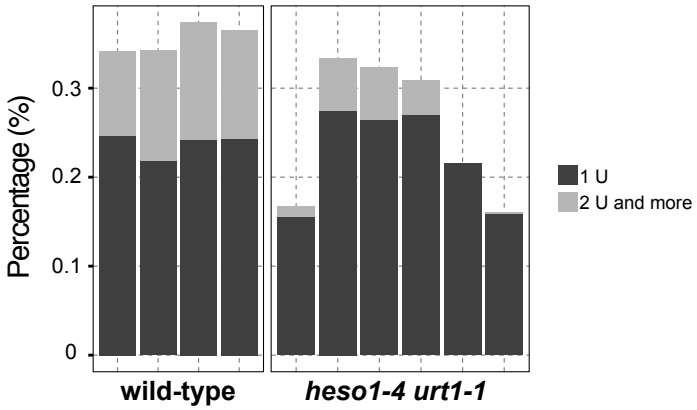


Figure 35: Percentage of uridylation on TYMV RNA in wild-type and *heso1-4 urt1-1*. The percentage has been calculated using the total number of reads for TYMV in each plant as denominator. Each bar represents one infected plant.

could be a competition between CCA-adding enzyme, aminoacyl-tRNA synthetase and TNTases for the access to the 3' extremity.

2. A-rich tails on viral RNAs could be involved in distinct mechanisms

Our results also revealed A-rich tails on viral RNAs, which have already been reported on plant viral RNAs (Li et al., 2014). Interestingly, we detected A-rich tails in two different viruses from the *Closteroviridae* family. More than 90% of GLRaV-2 RNA have onlyA and A-rich tails (Article Fig. 1D), that we detected at the expected mature 3' extremity (Article Fig. 1C). By contrast, 10% of GLRaV-1 RNA possess onlyA and A-rich tails (Fig. 18A), that were found on nibbled RNAs (Fig. 18C).

A-rich tails have been detected on prokaryotes, archaea and human RNAs (Bralley and Jones, 2002; Campos-Guillén et al., 2005; Mohanty and Kushner, 2000; Portnoy et al., 2005; Rott et al., 2003; Slomovic et al., 2006). For bacteria and archaea, the bacterial PNPase and the exosome are responsible for these additions (Mohanty and Kushner, 2000; Portnoy et al., 2005). In plants, A-rich tails have been observed in spinach chloroplasts and the chloroplastic PNPase has been identified as the polymerase activity (Lisitsky et al., 1996). Our team also identified A-rich tails on GFP reporter mRNA co-expressed in *N. benthamiana* as well as endogenous *PR2* mRNA (Scheer et al., 2021). A-rich tails have also been detected on plant viral RNAs (Li et al., 2014), but our team failed to confirm these results. So far, the cytoplasmic activity adding A-rich tails in the cytosol of plants remains unknown. In plants, a candidate is the RNA exosome which is able to degrade RNA from 3' to 5' but is also able to synthetize RNA tails (Sikorska et al., 2017). It would be worth to test whether GLRaV-1 RNA 3' A-rich tails are added by the RNA exosome, subsequently to 3' nibbling. Phylogenetic analyses have also revealed that Arabidopsis genome contain 19 genes that code for TNTases, which activities for some remain unknown to date (De Almeida et al., 2018a). So, A-rich tails could also be added by an unidentified cellular TNTase.

A possible explanation for the A-rich tails detected on 3' nibbled GLRaV-1 RNA is the presence of various viral isolates within the same plant. Indeed, GLRaV-1 originates from infected vineyards. It has been demonstrated in grapevine that infected vineyards can be infected by several viruses at the same time and that for one virus, different variants can co-exist within one plant (Alabi et al., 2011; Alliaume et al., 2018; Beuve et al., 2018; Fan et al., 2015; Komínek et al., 2005). At this stage of our investigation, we cannot exclude that the various 3' adenylated extremities that we mapped by 3'RACE-seq could correspond to GLRaV-1 variants.

Yet, an alternative explanation is that A-rich tails detected on GLRaV-1 RNA could be a mark for degradation. Indeed, polyadenylation has been reported to induce RNA decay in bacteria, in the nucleus, mitochondrion and chloroplast of plants, and in the nucleus and cytosol of metazoans (Briani et al., 2016; Lange and Gagliardi, 2022; Lange et al., 2009; Liudkovska and Dziembowski, 2021). In particular, oligoadenylation by Trf4-1 in drosophila was shown to stimulate mRNA degradation by the exosome (Harnisch et al., 2016). Because adenylated GLRaV-1 RNAs are nibbled, it is possible that they are degradation intermediates.

Figure 36: Model for the organization of PABP binding on the mRNA poly(A) tail. This model has been proposed in yeast with Pab1. For the Pab1 molecule proximal to the 3'UTR, RRMs1-3 bind ~22 As and RRM4 is associated to the 3'UTR. The distal Pab1 molecule bind ~28 As. Adapted from Webster *et al*., 2018.

3. Diverse patterns of viral RNA polyadenylation

Our 3'RACE-seq results on viral RNA polyadenylation in plants constitute a resource that could be useful for the scientific community interested in the multifunctional role of polyadenylation. It is the first extended analysis of viral RNA polyadenylation with a high-throughput sequencing method. Our results reveal an extreme diversity in the poly(A) tail size distribution for the viral RNA we analyzed (Fig. 19).

a. Protein binding could shape poly(A) tail size of viral RNAs

Our data suggest at least for some viral RNA that, as for mRNAs, viral poly(A) tails may be shaped by the binding of PABP. We can observe that several poly(A) tail size distributions display a peak at ~25 As (Fig. 19). In eukaryotes, initial work on PABP binding in yeast proposed a ~27 nucleotide footprint for PABP binding (Baer and Kornberg, 1983). A recent study in yeast has shown that the PABP bound closest to the 3'UTR binds to ~28 As or ~22 As, if the RRM1-3 bind to the poly(A) tail and RRM4 within the 3'UTR (Webster et al., 2018) (Fig. 36). The binding of PABPs on plant viral poly(A) tails could be possible for TuMV RNA. It has been reported that mRNA and protein levels of Arabidopsis PABP2/4/8 increase upon TuMV infection (Dufresne et al., 2008a). PABP is bound by TuMV VPg, causing its relocalization in viral factories (Beauchemin and Laliberté, 2007). Even if the role of cellular PABP in TuMV viral cycle remains unknown, this could illustrate the fact that we detect peaks at ~25 As and that PABP binding stabilizes viral RNAs in *planta*.

Interestingly, some viral poly(A) tails present a second peak, between 55 and 65 As (BNYVV, SLRSV RNAs; Fig. 19). The interval between the two peaks is ~30 As. A similar phasing pattern has been observed in *C. elegans* and in *in vitro* experiments (Lima et al., 2017; Webster et al., 2018). This type of distribution is expected to be found upon serial binding of PABP. In yeast, *in vitro* experiments have shown that two PABP molecules protect 50-55 adenosines (Webster et al., 2018). The poly(A) tail pattern for BNYVV and SLRSV RNAs could be in line with this PABP binding pattern. On PVX RNA, the distribution is shifted to the left, which could indicate that RRM4 of the first PABP binds to PVX RNA 3'UTR. The PABP binding pattern on viral RNAs could be confirmed by immunoprecipitating PABPs, as it was already done by our team to analyze PABP binding on mRNAs (Zuber et al., 2016).

PABPs are not the only proteins that are able to bind poly(A) tails. The LSm1-7 complex has been reported to preferentially bind oligoadenylated RNAs (<10 As) (Chowdhury et al., 2007) and could also explain some of the peaks that we observe in the poly(A) tail size distributions, such as for PVX RNA (Fig. 19).

In addition to proteins, the poly(A) tail can interact with elements within the sequence, resulting in RNA structures. The triple helix formed by the interaction with element for nuclear expression (ENE) and the poly(A) tail has been reported to prevent RNA decay (Conrad et al., 2006; Torabi et al., 2021a). ENE have been found in viral RNAs infecting mammals (Mitton-Fry et al., 2010; Torabi et al., 2021b). Similar interactions in plant viral RNAs could exist and would be worth studying.

b. The poly(A) tail size could be constrained by the capsid shape

Another hypothesis that could explain the diversity of poly(A) tail profiles is the shape and the form of the viral capsid. Theoretically speaking, viral RNAs encapsidated within icosahedral capsids are more constrained in size than those encapsidated within flexuous particles. Encapsidated RNAs in

icosahedral particles must not exceed a certain length. We did not analyze extensively this aspect for RNAs that display a narrow unimodal distribution such as CLRV RNAs. We did perform 3'RACE-seq on total RNA and RNA from purified virions for GFLV (Article Fig. 9), but for technical reasons our data do not allow to draw the robust poly(A) tail size distribution of encapsidated versus total RNA for GFLV RNAs.

In our data, whereas all viruses with a flexuous particle indeed possess long poly(A) tails (TuMV, BNYVV, PVX, GRSPaV, GPGV, GVB), we also observed long tails on viral RNAs for which the virion is icosahedral. For example, ArMV and SLRSV RNAs have long poly(A) tails, that could even exceed 90 As (Fig. 19) and the virion is icosahedral. By contrast, GRSPaV has flexuous viral particles and its RNA has a small proportion of poly(A) tails longer than 60 As (Fig. 19). Our results indicate that the shape and the form of the viral capsid are most likely not the only explanation for the diversity of poly(A) tail sizes on viral RNAs.

c. The poly(A) tail size could vary during the viral cycle

In addition, the various poly(A) tail sizes could correspond to different populations of viral RNAs, that are at a different step of the replication cycle. Short poly(A) tails could correspond to RNAs associated to polysomes. In *C. elegans* and mammalian cells, it has been demonstrated that mRNAs with a short poly(A) tail are translated more efficiently (Lima et al., 2017). For viral RNAs, an *in vitro* study has reported that a short poly(A) tail (<30 As) favors translation of wheat yellow mosaic virus RNA (*Potyviridae*) (Geng et al., 2019). Very recently, analyses on Sindbis virus RNA (*Alphavirus*) report that the minimum length of the poly(A) tail required for replication is 8 nucleotides and that RNAs tailed with less than 20 As are less efficiently translated, 20 As being the optimal length (Olsthoorn, 2022). Even though this manuscript is not peer-reviewed yet, it is in line with previous observations on poly(A) tail and translation efficiency.

In our analysis, the potyvirus TuMV RNA poly(A) tail size distribution displays indeed a peak around 20 As and high proportion of tails shorter than 30 As, as do BNYVV, GRGV and PVX RNAs (Fig. 19), which could correspond to a subpopulation being translated. So, it would be interesting to determine the poly(A) tail size of viral RNAs that are associated with translating polysomes. In collaboration with Rémy Merret (CNRS, Perpignan), our team tried to set up polysome purification to analyze polyadenylation on cellular mRNAs associated to polysomes. They purified polysomes either by sucrose gradient or by immunoprecipitating a FLAG-tagged RPL18. By comparing total RNA extracted from plants and RNA originating from polysome purification for both strategies, they observed technical RNA degradation in the samples originating from the polysome purification. So, the analysis of viral RNA polyadenylation on RNAs associated with polysomes requires testing and optimization, to try to minimize RNA degradation that could drive incorrect conclusions.

3. Distinct functions of URT1 and HESO1 in plants

1. Substrate and processivity differences on viral RNAs

Our analyses of viral RNA uridylation in *urt1-1* and *heso1-4* mutants have revealed differences in the RNA substrates of these TUTases. First, our data on TCV RNA point out the nucleotide preference: URT1 uridylates preferentially when the last 3' nucleotide is an A whereas HESO1 shows a preference for non-A 3' terminal nucleotides (Article Fig. 4C-E). This may partially explain that URT1 is the main TUTase uridylating TuMV RNA, since this RNA is polyadenylated (Article Fig. 6A). Our data also indicate that HESO1 uridylates a population of poly(A) tails that are shorter than those uridylated by URT1 (Article Fig. 6B, C).

It has already been reported that HESO1 and URT1 have substrate preferences. URT1 and HESO1 have been described to prefer a terminal uridine or adenosine, respectively. In the absence of the 3' terminal methylase HEN1, unmethylated miRNAs can be uridylated at their 3' extremity. miR158 ending with a uridine is preferentially uridylated by HESO1. By contrast, when it is truncated of one nucleotide, this miR ends with an adenosine and is preferentially uridylated by URT1 (Tu et al., 2015). In the case of miRNAs, the hypothesis that has been proposed is that URT1 mono-uridylates 1-nt trimmed miRNA, that is subsequently uridylated by HESO1 because of the substrate preference. Likewise, URT1 and HESO1 could cooperate to uridylate viral RNAs, with URT1 adding a few uridines to 3' terminal adenosine and HESO1 taking over to add longer U-tails once there is a 3' terminal uridine. For example, URT1 could uridylate TCV RNA when the last 3' nucleotide is an A and HESO1 could take over to add longer uridine extensions.

In line with this hypothesis, we also showed that the length of the uridine extensions varies between the two enzymes. Our 3'RACE-seq data on TuMV RNA has revealed an increased proportion of U-tails between 4 and 6 Us when URT1 is absent and this distribution is shifted toward 1-4 long U-tails when HESO1 is absent (Fig. 21A, B). This result implies that URT1 might mostly add short U-tails, by contrast to HESO1 which is proposed to be more processive. This difference of the number of U added by both TUTases has already been reported on miRNAs and RISC-cleaved fragments (Tu et al., 2015; Zhao et al., 2012; Zuber et al., 2018). For several unmethylated miRNAs, it has been shown that long U-tails are added by HESO1 and that URT1 adds 1 or 2 uridines (Tu et al., 2015; Zhao et al., 2012). U-tails longer than 2 Us on *SPL13* and *MYB33* 5' cleavage fragment disappear when HESO1 is absent but 1-2 Us extensions remain detected (Zuber et al., 2018). Hence, URT1 has been proposed to add 1-2 U-tails whereas HESO1 adds longer tails on miRNAs and RISC-cleaved fragments. Our results on viral RNAs tend to confirm these distinct processivities.

In addition, we observed that the proportion of U-rich tails on TCV RNA decrease significantly in *heso1-4* (Article Fig. 4A). This result suggests that HESO1 is responsible for the addition of long uridine extensions, that do not strictly contain U residues.

Our analysis on the activities of HESO1 and URT1 in uridylating viral RNAs shows that both TUTase display substrate preferences, different processivities and potentially distinct functions in viral RNA metabolism. HESO1-mediated uridylation could mediate the rapid decay of the RNA. In this case, the uridylated RNA would be difficult to detect by 3'RACE-seq in a wild-type background.

2. Similarities and differences between mRNA and TuMV RNA uridylation

Our data on TuMV RNA uridylation show similarities with mRNA uridylation in Arabidopsis. Indeed, TuMV RNA is uridylated at ~4% with 1-2 U-tails (Article Fig. 6A; Fig. 21A, B). Transcriptome-wide analysis on Arabidopsis has shown than most mRNAs are uridylated at 5% with 1-2 Us mainly ((De Almeida, 2019; Zuber et al., 2016) and unpublished data of C. de Almeida and H. Zuber). In addition, uridylation of TuMV RNA takes place on deadenylated RNA (Article Fig. 6B, C; Fig. 31), as it has been shown by our team for mRNAs in Arabidopsis (Scheer et al., 2021; Sement et al., 2013; Zuber et al., 2016). Therefore, it seems that deadenylation is also a prerequisite to TuMV RNA uridylation. Moreover, URT1 is the main TUTase involved in TuMV RNA uridylation (Article Fig. 6A), like it has been reported for mRNAs (Sement et al., 2013; Zuber et al., 2016).

Similarly to what was observed for mRNA (Zuber et al., 2016), our data show that uridylated poly(A) tails of TuMV RNA are shorter in *xrn4-3* than in wild-type (Fig. 31C). Yet, neither the level of uridylation nor the length of added uridine extensions detected on TuMV increased in *xrn4-3*, as compared to wild-type (Fig. 31C), contrasting with what our team previously observed for mRNAs (Zuber et al., 2016). The stabilization of very short poly(A) tails in *xrn4-3* suggests that XRN4 could be involved in the degradation of deadenylated TuMV RNAs. Whether uridylation on this population of short poly(A) tails is a signal for degradation remains to be determined. One possibility is that the population of RNAs with very short poly(A) tails, that are usually quickly degraded in wild-type, accumulate in the absence of XRN4, being thus accessible to the TUTases.

Recently, our team showed that URT1-mediated uridylation limits the accumulation of excessively deadenylated mRNAs, which triggers the formation of spurious siRNAs directed against endogenous mRNAs. Indeed, excessively deadenylated mRNAs accumulate in the absence of URT1 and, the absence of both URT1 and XRN4 triggers the biogenesis of illegitimate siRNAs targeting endogenous mRNAs. As a result, *urt1 xrn4* plants exhibit a strong morphological phenotype (Scheer et al., 2021). This phenotype was rescued by the mutation of DCL2 and DCL4, enzymes required for siRNA biogenesis (Ding and Voinnet, 2007). Our current model postulates that a key biological role of URT1-mediated uridylation is to prevent the accumulation of excessively deadenylated mRNAs, to avoid that those excessively deadenylated mRNAs accidentally enter the siRNA biogenesis pathway. Similarly, we could hypothesize that TuMV RNA uridylation by URT1 could hinder siRNA production, and therefore have a pro-viral role.

3. HESO1 overexpression negatively affects TuMV accumulation

For the three timepoints that I tested, the knock-out of *HESO1* did not affect TuMV accumulation (Fig. 23). Our results however show a significant lower viral accumulation in the plants overexpressing HESO1, especially in the Col-0 SOV (R705P) background (Fig. 25 and 27). TCV RNA is also uridylated by host TUTases (Article Fig. 4A). Preliminary experiments of TCV infections in plants overexpressing either one of the TUTases showed that the apparition of systemic symptoms seems to take longer in plants overexpressing HESO1 (Table 5).

Unpublished work in our team by a former PhD student, Caroline De Almeida, has shown that overexpressing HESO1 negatively impacts the steady-state level of a reporter GFP mRNA in a transient

expression system (De Almeida, 2019). So, HESO1-mediated uridylation could also result in a destabilization of the viral RNA. HESO1 could either uridylate the viral RNA or viral RNA fragments. In other organisms, 3'-5' decay by the SOV homolog DIS3L2 is enhanced by uridylation (Chang et al., 2013; Lubas et al., 2013; Malecki et al., 2013; Thomas et al., 2015; Ustianenko et al., 2013, 2016). Therefore, HESO1-mediated uridylation on TuMV RNA could enhance 3'-5' degradation by SOV.

A significant decrease in TuMV accumulation is observed in the line Col-0 SOV (R705P) HESO1-GFP (Fig. 25 and 27). This could be explained by the presence of a functional version of SOV but also by a higher expression level of HESO1 in this genetic background. The transgenic lines have been generated by floral dip and T-DNA insertion. The expression level of our protein of interest can vary depending on the number of insertions. So far, I analyzed one genetic line for this construction. To consolidate this observation, it is necessary to replicate the experiment with several other lines overexpressing HESO1 in Col-0 SOV (R705P) background, and to finely measure the level of HESO1 expression in the different transgenic lines to assess whether the effect, that is only observed in the Col-0 SOV (R705P) background, is due to the level of expression of HESO1 or to the presence of SOV.

The effect on TuMV accumulation could indeed be more pronounced in the presence of SOV, if SOV is involved in the decay of uridylated RNAs in the cytosol. The association between a TUTase and a 3'-5' exoribonuclease has been described in drosophila for Tailor and dmDis3l2. This complex has been named terminal RNA uridylation-mediated processing (TRUMP) (Reimão-Pinto et al., 2016). In this paper, the authors demonstrate that the TRUMP complex is involved in cytoplasmic decay. A similar cooperation could exist in Arabidopsis and explain why we observe a stronger effect of HESO1 on TuMV accumulation in the presence of SOV. Resolving SOV or HESO1 interactome could be a first step to understand whether SOV recruitment could be stimulated by uridylation and function in cytosolic decay.

However, indirect effects must also be considered. HESO1 overexpression could also impact host mRNAs involved in viral replication, thereby resulting in a less efficient viral replication. To investigate the potential transcriptome deregulation upon HESO1 overexpression and/or viral infection, it could be useful to perform RNA-seq on these plants. If deregulated transcripts are identified, then we could assess their uridylation status by 3'RACE-seq, or by DRS on a transcriptome-wide scale.

4. No substantial impact of URT1 on TuMV accumulation

Our data demonstrate that the absence of URT1 does not impact the global accumulation of TuMV (Fig. 23). I did not observe a robust impact of URT1 overexpression on TuMV accumulation (Fig. 25 and 27), even if more replicates are required to consolidate these observations.

In a recent work, our team has proposed a role of URT1-mediated uridylation in favoring the turnover of deadenylated mRNAs by recruiting decapping activators (Scheer et al., 2021). The fact that TuMV genome is not capped could explain that URT1 overexpression has no significant effect on TuMV accumulation. To investigate the impact of URT1 in stimulating the decapping, and subsequent viral RNA decay, it would be worth to test a capped viral RNA, like potexviruses PVX or the bamboo mosaic virus (BaMV).

Yet, URT1 has been reported to interact closely with decapping activator/translation inhibitor DCP5 (Scheer et al., 2021). So, URT1 could be involved in translation repression. In that case, translation

repression of TuMV RNA could impact the global accumulation of TuMV. Moreover, DCP1 has been described by our team to interact with the RNA endonuclease DNE1 (Schiaffini et al., 2022). So, recruitment of other RNA degradation factors to TuMV RNA via the URT1-decapping connection is not excluded.

5. Limitations of our experimental strategy

One limitation in quantifying viral accumulation, regardless of the method used, lies in the biological variability of the infection. When I plot the GFP signal intensity in wild-type plant by plant, especially at timepoints before 14 dpi, plants display variations (Fig. S2-S7). In wild-type plants, the variability is likely to be due to the plants and/or the inoculation efficiency and/or to the conditions of culture. This technical variability implies that many plants must be quantified, in order to be able to draw a conclusion that is based on a mutation rather than on plant-to-plant variability. In addition, quantifying the GFP signal to assess viral quantity is an indirect method. We cannot exclude that the GFP production is impaired but not the viral protein production. Nevertheless, it is a quick and relatively robust method to investigate viral infection, that is less time consuming than RT-qPCR. Moreover, it allows to follow the infection in the same plant over time, which is not possible when tissues have to be taken to perform RT-qPCR or ELISA assays.

My results are promising since they indicate a decreased accumulation of TuMV in transgenic lines overexpressing HESO1 in a Col-0 SOV (R705P) background, while no significant change was reported for URT1 overexpressing lines or for HESO1 overexpressing line in the *heso1-4* background. Yet, while I verified the expression of HESO1 and URT1 by western blot, further analyses are required to precisely quantify their level of expression. I do not know how much HESO1 is overexpressed, as compared to the endogenous protein. In addition, I did not verify whether the fusion protein is active. C-terminal fusion on URT1 has been reported not to affect URT1 function by our team (Scheer et al., 2021; Sement et al., 2013), but we did not check with a GFP tag. HESO1 fusion proteins have already been used in one study and it does not seem to impair its function, at least *in vitro* (Ren et al., 2014). Since we can see an effect on the global viral accumulation in HESO1 overexpressing plants, it suggests that the fusion is active and either has a direct effect by targeting the viral RNA, or an indirect effect by uridylating transcripts usually helping in viral replication.

6. Perspectives

Further work is required to elucidate by which mechanism the overexpression of HESO1 results in a global decrease of viral accumulation. One of the first experiment to perform could be to monitor the level of uridylation and the poly(A) tail distribution of the viral RNA in plants overexpressing TUTases. Overexpressing a TUTase could indeed result in a higher proportion of uridylated viral RNAs, that could be recognized by decay factors, thereby reducing the global viral accumulation. Of note, in this case, their fast turnover could prevent their detection.

In addition, the analysis of the global uridylation and polyadenylation profile upon TUTase overexpression in virus-infected plants could provide information on a potential deregulation of host transcripts. Host proteins like PABPC and translation initiation factors have been reported to be

important for TuMV replication. If they are destabilized by the overexpression of a TUTase, it could also result in a decrease of viral accumulation.

We did not observe significant effects when TUTases are absent or when URT1 is overexpressed. This lack of effects could lie in the choice of the virus. Indeed, TuMV accumulates at high levels in Arabidopsis, so the level of accumulation could reach such a high level that small accumulation differences are no more detectable. Moreover, the effects I observed when HESO1 is overexpressed mostly are early in systemic infection (8-10 dpi). To detect differences, we could analyze viral accumulation with viruses that accumulate less and/or at sooner timepoints. We could also harvest inoculated leaves instead of systematically infected leaves to investigate viral accumulation at the very beginning of the infection. In any case, monitoring viral accumulation in plants is challenging, given the biological variability due to the technique of infection, the culture conditions and the plant itself.

4. GFLV RNA peculiar uridylation: a new function for RNA uridylation?

1. What activity uridylates GFLV RNAs?

So far, we did not identify the factors involved in the mono-uridylation of GFLV RNAs. The single uridine could be added by another cellular enzyme, yet unidentified. Phylogeny analyses revealed that two candidate genes, *AT3G45750* and *AT3G45760*, cluster with the *HESO1* and *URT1* TUTases in Arabidopsis (De Almeida et al., 2018a). However, the nucleotidyltransferase activity of these proteins has not been characterized yet. Of note, these are adjacent genes and *AT3G45750* is constitutively expressed whereas *AT3G45760* expression is restricted to pollen. Preliminary data obtained in our team suggest that the protein encoded by *AT3G45750* localizes in the nucleus. In line with this observation, nuclear localization signal predictions in the sequence of these proteins identified potential NLSs in both of them. So, regarding GFLV RNA uridylation, it is unlikely that it is ensured by one of the candidate genes.

Another possibility lies in the uridine addition by the viral RNA-dependent RNA polymerase (RdRp). TNTase activity has been reported for several viral RdRps (Arnold et al., 1999; Behrens et al., 1996; Fullerton et al., 2007; Neufeld et al., 1994; Ranjith-Kumar et al., 2001; Rohayem et al., 2006; Smallwood and Moyer, 1993; Tomar et al., 2006; Tvarogová et al., 2019; Wang et al., 2013b; Wu et al., 2014). In particular, uridine preference has been described for the RdRp of the poliovirus, hepatitis C virus, Wuhan nodavirus and flock house virus (Arnold et al., 1999; Behrens et al., 1996; Wang et al., 2013b; Wu et al., 2014). So, it would be interesting to test the TNTase activity *in vitro* of GFLV RdRp, the 1^{E} protein.

2. Potential roles of GFLV RNA uridylation

a. Possible mechanisms involved in GFLV RNA uridylation

We cannot exclude that mono-uridylation of GFLV RNAs is linked to the genome replication process. To my knowledge, there is no molecular insight on the genome replication mechanism of GFLV.

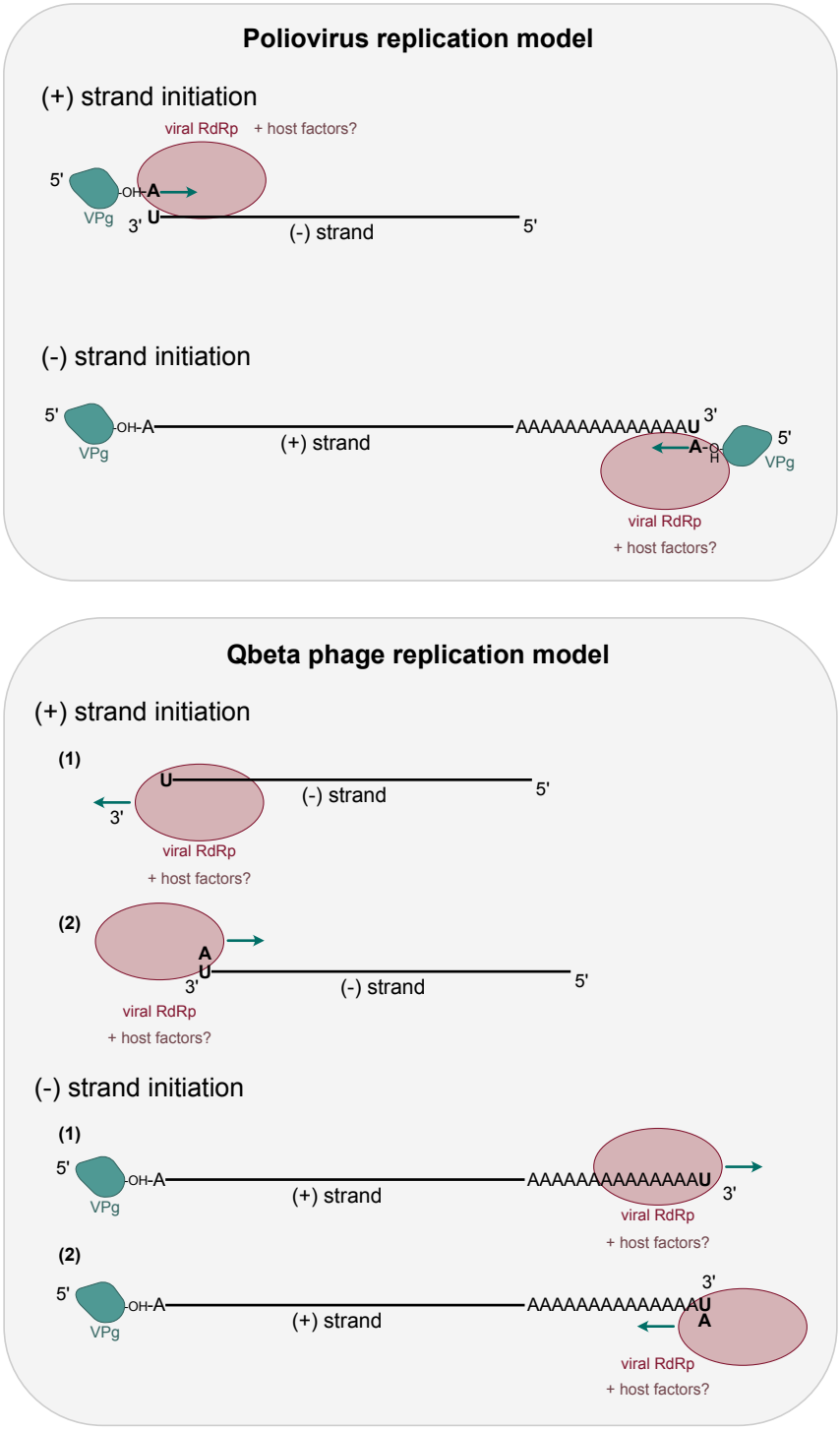


Figure 37: Models for GFLV RNA replication based on our data and published data on poliovirus and Qbeta phage replication. (A) Application of the poliovirus replication model to GFLV RNA. The VPg (green) is adenylated by the viral RdRp (red) and used to initiate the synthesis of (+) or (-) strand. Host factors could potentially be involved. In this model, both strands start with an A and end ith an U. **(B)** Application of the Qbeta phage replication model to GFLV RNA. (1) After the completion of the template strand synthesis, (2) the viral polymerase adds a 3' terminal U to the strand and starts to replicate it. Host factors could potentially be involved. In this model, both strands start with an A and end with an U.

Genome replication mechanism has been extensively studied for the poliovirus, member of the *Picornaviridae*, a family related to GFLV family, the *Secoviridae*. For polioviral RNAs, the replication relies on the poly(A) or poly(U) tail at the 3' extremity of the (+) or (-) strand, respectively, RNA structures, and the VPg (Steil and Barton, 2009; Steil et al., 2010). The initiation of the replication of the (-) strand is protein-primed by the VPg. First, the polioviral RNA circularizes through the interaction of RNA structures and proteins at the 5' extremity with the PABP bound to the poly(A) tail at the 3' extremity. A newly synthetized VPg is uridylylated at a tyrosine residue by the RdRp, resulting in VPg-pU-pU. It is then transferred to the 3' end of the poly(A) tail, probably in a complex with the viral RdRp. There, it can serve as primer to initiate (-) strand RNA synthesis. Hence, the first two 5' nucleotides are uridines in polioviral RNA (Paul and Wimmer, 2015).

For the GT isolate of GFLV in Arabidopsis, we have shown that the last 3' nucleotide of both the (+) and (-) strands is a uridine. On the (-) strand, this uridine is complementary to the 5' terminal adenosine of the (+) strand, which means that the terminal uridine on the minus strand is encoded by the viral genome. At the 5' extremity of the (-) strand for both RNAs, we detected a poly(U) tail (Article Fig. 7B and S5). So, both GFLV GT RNA strands have the same 3' terminal nucleotide but our result indicate that they have a different 5' terminal nucleotide. We have also shown that other isolates (B844, GHu, K30) have a 3' terminal uridine and a 5' terminal adenosine on the (+) strand (Article Fig. 7-9).

Considering the mechanism described for poliovirus, the conservation of the adenosine at the 5' extremity of the (+) strand and the presence of a 3' uridine on both strands, we could hypothesize that GFLV VPg could be adenylated, probably by the viral RdRp, thereby initiating the initiation of (+) strand. If it is the case, it would mean the both strands start with an adenosine (Fig. 37). Yet, our data indicate the presence of an oligo(U) at the 5' extremity of the (-) stand for both RNAs of GFLV GT in Arabidopsis and we did not detect any adenosine (Article Fig. 7B and S5). We cannot exclude that the absence of adenosine as first nucleotide of the (-) strand is due to technical limitations. Yet, our results show that a part of GFLV RNA poly(A) tail is encoded and replicated.

As we detected a single uridine at the 3' extremity of both RNA strands, we could also hypothesize that GFLV replication mechanism is similar to the Qß phage replication model. The Qß phage replicase, composed of a catalytic ß-subunit encoded by the phage genome and three host factors, terminates the replication by adding a single adenosine (Blumenthal and Carmichael, 1979). The GFLV polymerase would add a single uridine at the 3' extremity of the synthetized strand and then, switch template to synthetize the complementary strand (Fig. 37).

b. Does GFLV RNA mono-uridylation protect the 3' termini integrity?

Since the presence of one uridine has been reported to impede CAF1B deadenylase activity *in vitro* (Scheer et al., 2021), we could also hypothesize that GFLV RNA mono-uridylation represents a protecting feature, that is kept over the genome replication cycle, as it confers a selective advantage. Hence, GFLV RNA mono-uridylation could be a viral RNA feature characteristic for this virus, as are the A-rich tails on GLRaV-2 RNA.

Another argument in favor of this single uridine being a 3' terminal feature of GFLV RNAs is that we observed *in vivo* optimization of GFLV RNA 3' extremities. We analyzed GFLV RNA uridylation in *C.*

quinoa plants inoculated with infectious transcripts (isolate K30) (Article Fig. 9; Fig. 12). We also analyzed GFLV RNA uridylation in Arabidopsis inoculated with infected tissues from *C. quinoa* plants that were inoculated with infectious transcripts of the GT isolate (Article Fig. 7). Our collaborators in Colmar produced these transcripts by T7 RNA Polymerase *in vitro* transcription and the extremities of the viral transcripts are generated by restriction enzymes, after the sequence that codes for the poly(A) tail. The theoretical 3' extremities of the *in vitro* transcripts of the K30 isolate are composed of a poly(A) tail of 31 As and 22 As + 1G for RNA1 and RNA2, respectively, and additional nucleotides originating from the restriction site. For the GT isolate, the theoretical extremities are a poly(A) tail of 30 As + GC for RNA1 and a poly(A) tail of 22 As + 1G for RNA2. Our 3'RACE-seq data show that for both isolates, >80% of GFLV RNAs are uridylated with 1U (Article Fig. 7 and 9). In addition, we also showed for the GT isolate that the poly(A) tails of both RNAs mostly range from 22 to 65 As and that GFLV GT RNA poly(A) tail is uridylated, independently of its size (Article Fig. S4).

These results suggest that there is a strong selection pressure during GFLV replication in *planta*, resulting in the optimization of GFLV RNA 3' extremities. To gain insight on the importance of this i*n vivo* optimization, we want to produce *in vitro* transcripts that have poly(A) tails of different sizes, followed by a uridine or not. The aim is to test their efficiency to infect plants and to characterize the optimization of their extremities in *planta*. Our collaborators in Colmar are actually producing the *in vitro* transcripts.

5. First insights on uridylation and viral RNA cytosolic decay

1. No difference in global viral accumulation in the absence of decay factors

The contradictory results obtained by other groups have revealed biological variability regarding the role of XRN4 in viral RNA decay. Upon transient overexpression of XRN4 in *N. benthamiana*, TuMV accumulation is reduced (Li and Wang, 2018; Ma et al., 2019). In line with this observation, an increased level of TuMV was observed in a *xrn4-5* mutant (Ma et al., 2019). By contrast, one study reported that a mutation in *XRN4* leads to a decreased susceptibility to TuMV (Vogel et al., 2011). Some of the results presented in these papers may lack robustness, as they only quantified viral accumulation in a few plants. In addition, transient expression in *N. benthamiana* can also trigger plant immunity and induce effects that could be mistaken for virus-induced reactions. Our data on TuMV accumulation also reveals biological variability between the replicates, which does not allow to conclude on XRN4 impact on TuMV accumulation (Fig. 30).

I also investigated the impact of 3'-5' decay factors SOV, SKI2, and the catalytic subunit of the exosome RRP41 on TuMV accumulation. As for the absence of XRN4, I did not observe any significant change in the viral accumulation, except for both *rrp41*^{TYA} and *rrp41*^{TYA*} mutants (Fig. 30). Considering the biological variability illustrated by the individual points on the plot, this result needs to be repeated in more plants to be consolidated. Yet, this result is interesting and could give insights on the role of Exo9 in viral RNA decay *via* RRP41 catalytic activity.

The lack of differences in TuMV accumulation in our mutants as compared to wild-type is most likely explained by the redundancy of cytosolic decay pathways, as it has been reported in human cells

and Arabidopsis (Łabno et al., 2016; Sorenson et al., 2018). Global transcriptomic analyses have shown that decay factors have many overlapping substrates, so the absence of one could be compensated by another pathway. Inoculated plants mutated for several factors at the same time could allow to detect differences in viral susceptibility or resistance, as some decay factors are required for efficient viral multiplication like it has recently been shown for TuMV (Zuo et al., 2022).

As for some of the TUTases overexpressing lines, the similar levels of TuMV accumulation in wild-type and in decay mutants could be explained by the fact that TuMV accumulates at high levels, too high to detect any changes. So, we could assess viral accumulation at sooner timepoints or with viruses that accumulate less.

2. Accumulation of uridylated truncated RNAs in ski2-6

Even though we did not detect changes in the global viral accumulation in the absence of SKI2, our 3'RACE-seq data show that truncated uridylated TuMV RNA accumulate in *ski2-6* as compared to wild-type (Fig. 33). We detected uridylation on RNA fragments that have a 3' terminal uridine (Fig. 33C) and the U-tails are relatively long (Fig. 33D). These results could indicate that HESO1 is the main TUTase involved since it is more processive than URT1 and it prefers non-A nucleotides at the 3' end of the RNA substrate (Article Fig. 4; Fig. 21, (Tu et al., 2015; Zuber et al., 2018)). Our team and others already reported exosome-mediated decay of uridylated RNAs in the nucleus for ribosomal RNA, on which relatively long U-tails were detected (>10 Us) (Sikorski et al., 2015). In addition, unpublished work of a previous PhD student in our team, Natalia Sikorska, has shown *in vitro* that Exo9 prefers oligo(U) over oligo(A) (Sikorska, 2016).

The global viral accumulation does not change, indicating that TuMV RNA decay does not only rely on the action of SKI2 and the SKI complex. Since these fragments are stabilized in the absence of SKI2, the trimming activity could be performed by the cytosolic exosome, recruited to TuMV RNA by the SKI complex. These uridylated truncated fragments are most likely resulting from 3'-5' decay. TuMV RNA is deadenylated, further trimmed in the 3'UTR and uridylated. Successive rounds of trimming-uridylation result in fragments of different sizes. This promising result suggests a role of SKI2 and the cytosolic exosome in the decay of uridylated viral RNAs.

General conclusion

3'RACE-seq was an efficient strategy to analyze 3' tailing, in particular uridylation, of plant viral RNAs. Our work has contributed to establish a database of 3' tailing for representative members of phytovirus families. We analyzed viral RNAs ending with different type of extremities: tRNA-like structure or other types of RNA structures and poly(A) tail. Of note, due to our experimental strategy, we only captured non-aminoacylated TLS. We observed 3' uridylation on all the viral RNAs that we tested. Uridylation patterns display an extreme diversity: the proportion of uridylated viral RNA ranges from 0.3 to more than 80% and the uridine extension size varies from a strict mono-uridylation on GFLV RNAs to several uridines for the other viral RNAs.

Moreover, 3'RACE-seq has allowed to assess *in vivo* the 3' terminal addition of other nucleotides. We report 3' terminal adenylation on TYMV RNA TLS and 3' terminal mono-cytidylation on PCV RNA TLS. Our data also provide the first in-depth analysis of viral RNA polyadenylation profiles and indicate that viral RNA poly(A) tails have diverse size distributions.

We have identified two factors involved in viral RNA uridylation. Arabidopsis TUTases HESO1 and URT1 uridylate TuMV and TCV RNAs. TuMV RNA is mainly uridylated by URT1 and uridylation takes place on deadenylated TuMV RNAs. TCV RNA can be uridylated by both TUTases at similar levels but URT1 and HESO1 exhibit substrate preferences regarding the 3' recognized nucleotide: URT1 prefers a 3' terminal A whereas HESO1-mediated uridylation seems to be favored by non-A nucleotides. In addition, we also show that URT1 adds relatively short U-tails (1-3 Us) while HESO1 is able to synthetize longer tails, strictly composed of uridines but also uridines mixed with other nucleotides. Our results indicate that these TUTases have different substrate preferences and processivities on TuMV and TCV RNAs.

The primary role of uridylation described in eukaryotes is to induce RNA degradation, so we tested the impact of the absence or overexpression of Arabidopsis TUTases on TuMV accumulation. We did not observe any effect in the absence of the TUTases. Nevertheless, upon HESO1 overexpression, we were able to show that TuMV accumulation is reduced when a functional version of SOV, the 3'-5' exoribonuclease DIS3L2 homolog, is present. This result suggests that HESO1-mediated uridylation is related to RNA destabilization. The mechanisms that are in place have not been identified yet, but could be of potential interest to control viral accumulation in plants.

In addition, in the absence of the RNA helicase SKI2, a cofactor of the cytosolic exosome, we observed the accumulation of trimmed TuMV RNA tailed with long uridine extensions. The stabilization of uridylated trimmed TuMV RNA in the absence of SKI2 could indicate that TuMV RNA uridylation favors its decay via the 3'-5' decay pathway by the SKI complex and the exosome.

Our 3'RACE-seq analysis has also revealed a singular uridylation pattern on GFLV RNAs which has, to our knowledge, never been observed on an RNA in plants. We have demonstrated that Arabidopsis TUTases are not involved in the maintenance of this pattern. Although GFLV RNA uridylation activity has not been identified yet, it is possible that this uridine could be added by the viral

RNA polymerase. This role remains to be tested. To date, the role of uridylation on GFLV RNA also remains unknown, but it could reveal a new molecular mechanism associated to RNA uridylation.

Overall, our results highlight the diversity of viral RNA uridylation profiles in plants. This extreme diversity suggests that potentially, several mechanisms related to plant viral RNA uridylation are in place.

Material and Methods

1. Material

1. Plant material

a. Plant growth conditions

The *Arabidopsis thaliana* plants used in this work are of Columbia accession (Col-0). For inoculation experiments, plants are grown on soil with 12 h light / 12 h darkness cycles in a neon-lit chamber (at 21/18 °C). For selecting TUTases overexpressing lines, plants are grown on selective medium (MS 255 0.44% w/v; sucrose 1% w/v; agarose 0.8% w/v; pH 5.7 adjusted with KOH 1M; hygromycin at 25 ug/mL or glufosinate at 10 ug/mL) and resistant plants are transferred into soil after 10-12 days to grow at 21/18 °C for 16 hours day / 8 hours night light conditions (long day conditions).

Nicotiana benthamiana (wild-type and B2), *Nicotiana clevelandii*, *Spinacea oleracea*, *Chenopodium quinoa and Brassica napus* plants used for agroinfiltration or viral inoculation are grown on soil with 16 h light / 8 h darkness cycles in a greenhouse (at 21/18 °C). The *N. benthamiana 35S:B2:*GFP line used for viral inoculation was previously described (Monsion et al., 2018).

Infected grapevines were either from INRAE collection in greenhouse (vines in individual pots and grown to obtain two shoots of 180 cm, natural light conditions and no heat in a greenhouse) or from vineyard.

b. Mutants

T-DNA insertion lines of URT1 (*AT2G45620*, *urt1-1* SALK_087647C), HESO1 (*AT2G39740*, *heso1-4* GK_369H06_017072), SKI2 (*AT3G49690*, *ski2-6* SALK_122393) and of XRN4 (*AT1G54490.1*, *xrn4-3* SALK_014209) were used. The *urt1-1* and *xrn4-3* alleles were previously characterized (Gazzani et al., 2004; Sement et al., 2013). The HESO1 and the SKI2 mutants were identified from the GABI-KAT and SALK collections (Alonso et al., 2003; Rosso et al., 2003). The double mutant *heso1-4 urt1-1* was obtained by crossing the simple mutants and kindly provided by P. Brodersen's group.

The exosome mutants *rrp41* RRP41^{Pi} (*rrp41*^{TYA} in this manuscript) and *rrp41* RRP41$^{Pi\text{-}Cat}$ (*rrp41*^{TYA*} in this manuscript) were obtained by agrobacterium-mediated transformation of heterozygous RRP41/*rrp41* (SALK_112819) plants (Sikorska et al., 2017).

The Col-0 SOV (R705P) mutant was ordered from the Nottingham Arabidopsis Stock Center (NASC, ID: N66970, SOVCol-0R705P). These plants were generated by agrobacterium-mediated transformation of Col-0 with a binary vector expressing SOV (*AT1G77680.1*) mutagenized to code for proline at position 705 instead of arginine, corresponding the version of SOV found in the Landsberg accession (Zhang et al., 2010). In their work, Zhang and colleagues propose that the arginine found at position 705 in SOV$^{Col\text{-}0}$ makes the protein inactive, compared to SOV^{Ler}. After receiving the seeds, homozygous plants for the insert were selected by genotyping. To investigate the potential role of SOV

in viral RNA degradation, we worked with Col-0 and Col-0 SOV (R705P) plants to investigate the potential difference of activity.

URT1 and HESO1 overexpressing plants were obtained by agrobacterium-mediated transformation (Bernhardt et al., 2012) of homozygous *urt1-1* and *heso1-4* plants, respectively, with binary vectors expressing URT1 or HESO1 (see Methods, Chapter 2.2).

2. Virus isolates

The virus isolates used in this work are listed in Supplementary Data 1b, available from: https://data.mendeley.com/datasets/k3njd5yfj3/draft?a=bd2e49d5-ac42-4b84-befb-23e6df7e2011

3. Bacterial strains

Chemically competent *Escherichia coli* TOP10 cells were used for plasmid amplification.

Chemically competent *Agrobacterium tumefaciens* GV3101 was used to transform plants. This strain has rifampicin resistance in its genome and the Ti plasmid pmp90 with gentamycin resistance. The Ti plasmid is a helper plasmid in which the T-DNA region sequences are deleted and transformation with a binary vector containing the missing T-region produces a functional T-DNA binary system. It can then be used to transfer genetic material into a host plant's genome.

4. Vectors

pDONR™Zeo and **pDONR™207 are** donor Gateway® vectors with attP1 and attP2 sites, a zeocin and a gentamycin resistance marker, respectively, and the *ccdb* gene, classically used for selecting Gateway® cloning products.

pGWB17 is destination Gateway® vector with attR1 and attR2 sites, a spectinomycin resistance marker for bacterial expression and both kanamycin and hygromycin resistance markers for plant expression. This plasmid is used to express sequences of interest fused to four myc units in C-terminal. The construct is expressed under the control of the 35S CaMV promoter and nopaline synthase terminator. In this study, pGWB17 has been used to generate plants overexpressing HESO1-4xmyc constructs, its catalytic mutant.

pUBC-GFP and pUBC-RFP are destination Gateway® vector with attR1 and attR2 sites, a spectinomycin resistance marker for bacterial expression and glufosinate resistance marker for plant expression. This plasmid is used to express sequences of interest fused to the green fluorescent protein (GFP) or the red fluorescent protein (RFP), respectively, in C-terminal. The construct is expressed under the control of the ubiquitin-10 promoter (AT4G05320.2). In this study, **pUBC-GFP and pUBC-RFP** have been used to generate plants overexpressing HESO1-GFP, URT1-GFP and HESO1-RFP constructs.

5. Antibodies

The **@myc** antibody is a mouse monoclonal antibody produced by Roche. It is used at a dilution of 1/10 000. It is recognized with secondary goat antibodies coupled to horseradish peroxidase (@GAM).

The **@URT1** antibody is a rabbit serum produced by our team. It is used at a dilution of 1/10 000. It is recognized with secondary goat antibodies coupled to horseradish peroxidase (@GAR).

The **@GFP** antibody is a rabbit serum produced by D. Gilmer's lab. It is used at a dilution of 1/50 000. It is recognized with secondary goat antibodies coupled to horseradish peroxidase (@GAR).

The **@GFP-HRP** antibody binds to GFP and is coupled to horseradish peroxidase (Miltenyi Biotech).

6. Primers

a. Primers used for 3' and 5'RACE-seq

Target	Primer sequence (5'-3')	Description
AIMV RNA1	CATGAAGCTCAGGGTAAGAC	3'RACE-seq PCR1 forward primer
AIMV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGTT GCC TTG TCG AGA CAC AAG	3'RACE-seq PCR2 forward primer
AIMV RNA2	CACCATGTTATCCGATGCGTTG	3'RACE-seq PCR1 forward primer
AIMV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGAGTTTATACCTAGAAGCTGC	3'RACE-seq PCR2 forward primer
AIMV RNA3	CGCAGTTTGATTACTGTGG	3'RACE-seq PCR1 forward primer
AIMV RNA3	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGATCGTCATTGATGTACCCC	3'RACE-seq PCR2 forward primer
ArMV RNA1	GCCCTCTATTCCCATGTTGATG	3'RACE-seq PCR1 forward primer
ArMV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCCTCCACACTTCAACTTTTGGG	3'RACE-seq PCR2 forward primer
ArMV RNA2	CATGGGGTGTGAGACTGAAC	3'RACE-seq PCR1 forward primer
ArMV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGTAACTTTGCTGGTGCCAATCC	3'RACE-seq PCR2 forward primer
BBWV-1 RNA1	TGTCATAAGGAAAGGCTGTGC	3'RACE-seq PCR1 forward primer
BBWV-1 RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCTCCGTCATACCAGCATATG	3'RACE-seq PCR2 forward primer
BBWV-1 RNA2	CAAAAGGACAAGCTCACATG	3'RACE-seq PCR1 forward primer
BBWV-1 RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCAGGGTTTTCTATCCGATGGCTG	3'RACE-seq PCR2 forward primer
BNYVV RNA1	GCTGTGTCTATTGGCGAC	3'RACE-seq PCR1 forward primer
BNYVV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCAC AGG CCT CCT ATC TTG	3'RACE-seq PCR2 forward primer
BNYVV RNA2	GTGTTTGCGGTGATTGTTGTG	3'RACE-seq PCR1 forward primer
BNYVV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGTCGATCCTGAGGTGTAAC	3'RACE-seq PCR2 forward primer
BNYVV RNA3	CGAAGGTTTAACTGTGAGCC	3'RACE-seq PCR1 forward primer

BNYVV RNA3	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCCTAACAGGCTCGTTCAC	3'RACE-seq PCR2 forward primer
BNYVV RNA4	GGCGAAGGTTTTGTGATGTTC	3'RACE-seq PCR1 forward primer
BNYVV RNA4	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGAGTCTCGTACTTTTGCACG	3'RACE-seq PCR2 forward primer
CLRV RNA1	CCTAGATCTTATTGCCGTGTACTCC	3'RACE-seq PCR1 forward primer
CLRV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCTTACTTGGTATATGAGCCGGG	3'RACE-seq PCR2 forward primer
CLRV RNA2	TGGTTTCAGGTTTCGGGGTC	3'RACE-seq PCR1 forward primer
CLRV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCTTACTTGGTATATGAGCCGGG	3'RACE-seq PCR2 forward primer
CMV RNA1	GCAAACCGTCTGAAGTCAC	3'RACE-seq PCR1 forward primer
CMV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGGTTGTCCATCCAGCTAAC	3'RACE-seq PCR2 forward primer
CMV RNA2	CGGACGAGCTAGTACTTCATG	3'RACE-seq PCR1 forward primer
CMV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCGGAAGGTGCTTTCTGAAAC	3'RACE-seq PCR2 forward primer
CMV RNA3	CGGACGAGCTAGTACTTCATG	3'RACE-seq PCR1 forward primer
CMV RNA3	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGTTCCCAGAATCCTCCCTCC	3'RACE-seq PCR2 forward primer
CNDV	CGGTTTGAAAGAGCGCATACTG	3'RACE-seq PCR1 forward primer
CNDV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGGCTTATCCTAGGGGTGATTAGT	3'RACE-seq PCR2 forward primer
CPMV RNA1	TTCATTGCCTGATGGGTGTC	3'RACE-seq PCR1 forward primer
CPMV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGTGATATTCTGTGCTCGTG	3'RACE-seq PCR2 forward primer
CPMV RNA2	CCAATCCCAGACAAATACAGC	3'RACE-seq PCR1 forward primer
CPMV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCTCCAAGCGTAGTGTTATGG	3'RACE-seq PCR2 forward primer
GFLV B844 RNA1	AGTGCTGCCAAAGGTGTGTAG	3'RACE-seq PCR1 forward primer
GFLV B844 RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCATCACTCGTGCTCTAGGAAGTAG	3'RACE-seq PCR2 forward primer
GFLV B844 RNA2a and RNA2b	CTCCCTTTATAGCCGATGGATG	3'RACE-seq PCR1 forward primer
GFLV B844 RNA2a and RNA2b	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGCTTTAACTTGTTTACTGC	3'RACE-seq PCR2 forward primer
GFLV Ghu RNA1	ATYCTTGGTACTGGGATAACC	3'RACE-seq PCR1 forward primer
GFLV Ghu RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCAAGYTTAAATAACCCAGTTTC	3'RACE-seq PCR2 forward primer
GFLV Ghu RNA2	GCCTGGATGGACCTTCTCAG	3'RACE-seq PCR1 forward primer

GFLV Ghu RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCACTCCGTGTTTTGTGCAAGC	3'RACE-seq PCR2 forward primer
GFLV GT RNA1 (-)	5'-CCGCATTTGAAAAGGACTTCCC-3'	3'RACE-seq PCR1 forward primer on minus strand, RNA1
GFLV GT RNA1 (-)	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCAGCGAGTGCAACGAAAAGGT	3'RACE-seq PCR2 forward primer on minus strand, RNA1
GFLV GT RNA1 (+)	GTCGCCAAAACTGTGAGTAGT	3'RACE-seq PCR1 forward primer on plus strand, RNA1
GFLV GT RNA1 (+)	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTCAAGCTCTTTAAGGCGTCTG	3'RACE-seq PCR2 forward primer on plus strand, RNA1
GFLV GT RNA2 (-)	CCCAAGATGGGGATTCTTTCCA	3'RACE-seq PCR1 forward primer on minus strand, RNA2
GFLV GT RNA2 (-)	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCGCTTTGGTACCGCTCTTAAC	3'RACE-seq PCR2 forward primer on minus strand, RNA2
GFLV GT RNA2 (+)	CAAAACTGGTTGGTGACAAGG	3'RACE-seq PCR1 forward primer on plus strand, RNA2
GFLV GT RNA2 (+)	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGTACTCCGTGTTTTGTGCAAGC	3'RACE-seq PCR2 forward primer on plus strand, RNA2
GLRaV-1	CACTTTTCGCTTCTTCGGAG	3'RACE-seq PCR1 forward primer
GLRaV-1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCGATGTCCGTCTATGCGAGA	3'RACE-seq PCR2 forward primer
GLRaV-2	RTTCATCACGCGAAAGTACG	3'RACE-seq PCR1 forward primer
GLRaV-2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTCGAACGTGGGTRTATCTAC	3'RACE-seq PCR2 forward primer
GPGV	GCCGGTCGTGAAGGAATACA	3'RACE-seq PCR1 forward primer
GPGV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCACCCCAAATCAAAAGTCGGTG	3'RACE-seq PCR2 forward primer
GRGV	TACGCCACTCTCCAATCTCT	3'RACE-seq PCR1 forward primer
GRGV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCACAAAAACCCTGAAGCCCCA	3'RACE-seq PCR2 forward primer
GRSPaV	ATCCATYTGGAGACGGGRAT	3'RACE-seq PCR1 forward primer
GRSPaV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCTTGGAGARATTAGTGGYGG	3'RACE-seq PCR2 forward primer
GRVFV	TCAGCCTACTCTTCCTCTCG	3'RACE-seq PCR1 forward primer
GRVFV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGTGGCATGGTCAGCGTCTCT	3'RACE-seq PCR2 forward primer
GVB	TGGAGTGCAATGGAGGAAAG	3'RACE-seq PCR1 forward primer
GVB	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGCAAGGTTAGGCATGTTCGC	3'RACE-seq PCR2 forward primer

PCV RNA1	CGAGCAATTCTGTGGAGAG	3'RACE-seq PCR1 forward primer
PCV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCTGAGAAGACTGGAGAGGCA	3'RACE-seq PCR2 forward primer
PCV RNA2	CCTCATGTGGATGCAGTA	3'RACE-seq PCR1 forward primer
PCV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGTGTGCCTATTGATCCAAAGG	3'RACE-seq PCR2 forward primer
PVX	GCAAGCACAAGGTTTCAAGCC	3'RACE-seq PCR1 forward primer
PVX	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCGCTGCATTCGACTTCTTCA	3'RACE-seq PCR2 forward primer
RpRSV RNA1	TCGGTAATCAGACAAGTGCC	3'RACE-seq PCR1 forward primer
RpRSV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCGGTATATGGCAACTCTCTG	3'RACE-seq PCR2 forward primer
RpRSV RNA2	GCCAACTAAAGCTGTACCATG	3'RACE-seq PCR1 forward primer
RpRSV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCAACTTGTAGGTTCAGCGG	3'RACE-seq PCR2 forward primer
SLRSV RNA1	CCGGTGAACTTCGTGTGTCA	3'RACE-seq PCR1 forward primer
SLRSV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCACACGCTTGTAGACCGTACG	3'RACE-seq PCR2 forward primer
SLRSV RNA2	TGTCCATGTGTTGAGGCTGA	3'RACE-seq PCR1 forward primer
SLRSV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCCACACGCTTGTAGACCGTA	3'RACE-seq PCR2 forward primer
SoMV	GGCACGCACTGAACTTAG	3'RACE-seq PCR1 forward primer
SoMV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGGAAGCAATCGCCACTATTG	3'RACE-seq PCR2 forward primer
TBSV	GACAACGCATACAGGTTACTC	3'RACE-seq PCR1 forward primer
TBSV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATC. GTTATGATGACGAGTCGGTC	3'RACE-seq PCR2 forward primer
TBRV RNA1	GCAGCAAAGGAGAGGAAGGT	3'RACE-seq PCR1 forward primer
TBRV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGTGCCCCCATCATCTTGCTA	3'RACE-seq PCR2 forward primer
TBRV RNA2	TGGTTCTTTTGCTGGACCCA	3'RACE-seq PCR1 forward primer
TBRV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCACCAGTTTGACCGTCAGCAT	3'RACE-seq PCR2 forward primer
TCV	CAATGGGCAGGAGTGAAGGTA	3'RACE-seq PCR1 forward primer
TCV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCAGCACTCAGAATTTAGTACGG	3'RACE-seq PCR2 forward primer

TMV	TTGAAAATCAGGCGAACCCCAC	3'RACE-seq PCR1 forward primer
TMV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGACTGCCGAAACGTTAGATGCT	3'RACE-seq PCR2 forward primer
TRV RNA1	GTTGTGGCCGTAGTCACCTTG	3'RACE-seq PCR1 forward primer
TRV RNA1	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCATGTCTGCGACAGCTAAAAAG	3'RACE-seq PCR2 forward primer
TRV RNA2	GTAGATCCGTGTCTGAAG	3'RACE-seq PCR1 forward primer
TRV RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGACATTCTCGACTGATCTTG	3'RACE-seq PCR2 forward primer
TRSV RNA1	GGGCTAGYTCTTTGAGGAGYTAC	3'RACE-seq PCR1 forward primer
TRSV RNA2	GGYGCCTTGGGTAAGCAAAAC ; GGTGTCTTGGGTAAGCAAAAC ; GGTGCTTTGGGTAAGCAAAAC	3'RACE-seq PCR1 forward primer
TRSV RNA1&RNA2	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCAGCTGCTTGTGTAAATGAGC	3'RACE-seq PCR2 forward primer
TuMV	CCAGATGAAAGCAGCAGCACT	3'RACE-seq PCR1 forward primer
TuMV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCAGGACGTTAATCGGAACATGC	3'RACE-seq PCR2 forward primer
TuYV	GGAAGCGCGTTAATACCACG	3'RACE-seq PCR1 forward primer
TuYV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGTACTTAACTGGCTCGTTG	3'RACE-seq PCR2 forward primer
TYMV	CACCAAGACCTATGGTGG	3'RACE-seq PCR1 forward primer
TYMV	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGATTCGATTCAGTACCTTGAC	3'RACE-seq PCR2 forward primer
GFLV GT_R1_minus_1.2	GCCACCAGCTCAGATACTAG	forward primer for 5'RACE-seq PCR1 on minus strand, RNA1
GFLV GT_R1_minus_2.2	CATCTTAAGACTCTGAGGCGTG	forward primer for 5'RACE-seq PCR2 on minus strand, RNA1
GFLV GT_R1_minus_2.2 _Illu	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCATCTTAAGACTCTGAGGCGTG	forward primer for 5'RACE-seq PCR3 on minus strand, RNA1
GFLV GT_R2_minus_1.2	GGGTTGCACAGCGGCATATTGG	forward primer for 5'RACE-seq PCR1 on minus strand, RNA2
GFLV GT_R2_minus_2.2	GGCAAGAGTAGCGGGAGCGT	forward primer for 5'RACE-seq PCR2 on minus strand, RNA2
GFLV GT_R2_minus_2.2 _Illu	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGGCAAGAGTAGCGGGAGCGT	forward primer for 5'RACE-seq PCR3 on minus strand, RNA2
GFLV GT RNA1 plus 1	CCGCATTTGAAAAGGACTTCCC	forward primer for 5'RACE-seq PCR1 on plus strand, RNA1

GFLV GT RNA1 plus 2	CAGCGAGTGCAACGAAAAGGT	forward primer for 5'RACE-seq PCR2 on plus strand, RNA1
GFLV GT_RNA1_plus_2 _Illu	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCCAGCGAGTGCAACGAAAAGGT	forward primer for 5'RACE-seq PCR3 on plus strand, RNA1
GFLV GT RNA2 plus 1	CCAAGATGGGGATTCTTTCCA	forward primer for 5'RACE-seq PCR1 on plus strand, RNA2
GFLV GT RNA2 plus 2	GCTTTGGTACCGCTCTTAAC	forward primer for 5'RACE-seq PCR2 on plus strand, RNA2
GFLV GT_RNA2_plus_2 _Illu	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGACGATCGCTTTGGTACCGCTCTTAAC	forward primer for 5'RACE-seq PCR3 on plus strand, RNA2
3' adapter	/5rApp/CTGACNNNNNNNNNNNNNNNTGGAATTCTCGGGTGC CAAGGC/3ddC/	3' adapter for 3'RACE-seq
3'-RT	GCCTTGGCACCCGAGAA	primer for RT 3'RACE-seq
3'RACE_reverse_ PCR1	CCTTGGCACCCGAGAATTCCA	reverse primer for 3'RACE-seq PCR1
3'RACE_reverse_ PCR2	True Seq Illumina	reverse primer for 3'RACE-seq PCR2
5'RACE_reverse_ PCR1	CCTTGGCACCCGAGAATTCCANNNNNNNNCTGACGGGIIGGG IIGGGIIG	reverse primer for 5'RACE-seq PCR1
5'RACE_reverse_ PCR2	CCTTGGCACCCGAGAATTCCA	reverse primer for 5'RACE-seq PCR2
5'RACE_reverse_ PCR3	True Seq Illumina	reverse primer for 5'RACE-seq PCR3
Poly A Spike_N	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNNN NNNNNNCTGACGAGCTACTGTTGGAATTCTCGGGTGCCA	Spike-in
Poly A Spike_8	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAA AAAAAACTGACGAGCTACTGTTGGAATTCTCGGGTGCCA	
Poly A Spike_16	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAA AAAAAAAAAAAAAACTGACGAGCTACTGTTGGAATTCTCGGG TGCCA	
Poly A Spike_32	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGACGAGCTACT GTTGGAATTCTCGGGTGCCA	
Poly A Spike_50	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAACTGACGAGCTACTGTTGGAATTCTCGGGTGCCA	
Poly A Spike_64	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAACTGACGAGCTACTGTTGGAATTCTC GGGTGCCA	
Poly A Spike_80	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGACGAGC TACTGTTGGAATTCTCGGGTGCCA	

b. Primers used for RT-qPCR

Target	Primer sequence (5'-3')	Description
GT_RNA1_qPCR_fw2	GATGTACGCCTGAAGGGAACA	qPCR fwd primer
GT_RNA1_qPCR_rev2	CGTCTGAACCATATCACCTGC	qPCR rev primer
GT_RNA2_qPCR_fw1	GGTGTCCAGTATGAAAAGTGG	qPCR fwd primer
GT_RNA2_qPCR_rev1	ACTTATGGTGGATAAGCC	qPCR rev primer
TuMV_qPCR_fwd	TCACCATCAAAACACCAGCGA	qPCR fwd primer
TuMV_qPCR_rev	CCATCCACAATCCACCTCCAAG	qPCR rev primer
ACT2q-F	GCACCCTGTTCTTCTTACCG	qPCR fwd primer
ACT2q-R	AACCCTCGTAGATTGGCACA	qPCR rev primer

c. Primers used for Gateway cloning

Target		Primer sequence (5'-3')	Description
AT2G3974	HESO1 mut FW	CACACGATGGGGAGCCTTAGCTATCTCTG	forward primer for site-directed mutagenesis of HESO1 catalytic site (D66A+D68A)
AT2G3974	HESO1 mut RV	CAGAGATAGCTAAGGCTCCCCATCGTGTG	reverse primer for site-directed mutagenesis of HESO1 catalytic site (D66A+D68A)
AT2G3974	GW FW HESO1	GGGGACAAGTTTGTACAAAAAAGCAG GCTTCATGAGTAGAAACCCTTTC	Gateway cloning forward primer
AT2G3974	GW RV HESO1 noSTOP	GGGGACCACTTTGTACAAGAAAGCTGG GTCCTGCTCATGTCTCGGTCTC	Gateway cloning reverse primer
AT2G45620	URT1-fw	GGGGACAAGTTTGTACAAAAAAGCAGG CTCATTGTAAGATCTTCGACTTGGC	Gateway cloning forward primer
AT2G45620	URT1-noStop-rev	GGGGACCACTTTGTACAAGAAAGCTGGG TAGTTGTGGCCTTGTCCATTATTAT	Gateway cloning reverse primer

d. Primers used for genotyping T-DNA insertion lines

Mutant	Primer name	Primer sequence (5'-3')
urt1-1 SALK_087647C	wt_fwd	ACCTGTGGATGTCATTTCGAC
urt1-1 SALK_087647C	wt_rev	GTTTCCTCAATTTCCGCCTAG
urt1-1 SALK_087647C	m_fwd	CCGTCAATAATCCTTTTCCTCCG
heso1-4 GK-369H06	heso1-4_fwd	CTTTACTCGGACATTTACTACGAGC
heso1-4 GK-369H06	wt_rev	CACCGAACTCTTGGGCCTTC

xrn4-3 SALK_014209	xrn4-3_fwd	TCCCATGAGAGCCATGCATTC
xrn4-3 SALK_014209	wt_rev	ACCATCCTGCAGGTCCAAGAA
ski2-6 SALK_122393C	ski2-6_fwd	CGCCAAGCTTTTTGTAGTCTC
ski2-6 SALK_122393C	wt_rev	CCGTAGGTGAACCTGCGGAGG
GABI T-DNA insertion line	GABI_rev	ATAATAACGCTGCGGACATCTACATTTT
SALK T-DNA insertion line	SALK_rev	ATTTTGCCGATTTCGGAAC
AT1G77680 SOV (R705P)	SOV_insert_fwd	TTTGGTGAGTCCTTGTGAAGTG
AT1G77680 SOV (R705P)	SOV_insert_rev	GCTCTAGCATTCGCCATTCAG
AT1G77680 SOV (R705P)	SOV_mut_fwd	GACAGCTACAGATCTCAATGATGC
AT1G77680 SOV (R705P)	SOV_mut_rev	GAGGTTTCCAGTGATTTTCTCC

2. Methods

1. Gateway Cloning

The Gateway cloning strategy is based on intrinsic site-specific recombination properties of the lambda bacteriophage. First, the sequence of interest is inserted into a Gateway® donor plasmid using the BP Clonase® II enzyme (Invitrogen). This enzyme assures recombination of the *attB* sites of the insert amplified by PCR with the *attP* sites of the donor plasmid (p207 in this study). The entry vector containing the inserted sequence of interest flanked by *attL* sites is obtained and amplified in bacteria. Then, the recombination of the *attL* sites of the entry plasmid with the *attR* sites of the Gateway destination vector is performed by the LR Clonase® II enzyme (Invitrogen). The obtained vector can further be used for bacterial or plant expression.

a. PCR amplification of the sequence of interest (C. De Almeida)

The *HESO1* cloning was initiated by Caroline De Almeida, a former PhD student in our team. The coding sequence of *HESO1* (AT2G3974) with no stop codon is amplified by polymerase chain reaction (PCR) using the Phusion DNA Polymerase (Thermo Scientific) from complementary DNA (cDNA). The *attB1* and *attB2* sequences are introduced during the PCR. The reaction is performed in 20 µL containing 0.5 µM forward and reverse primers specific to the sequence of interest and that contain the *attB* cloning site (1. Material 6.c.), 1X Phusion Master Mix (Thermo Scientific) and 30 ng of cDNA. The program of amplification is the following: 2 min denaturation at 98 °C, 35 cycles of 30 s denaturation at 98 °C, 45 s at the annealing temperature of the primers, 90 s elongation at 72 °C, a final elongation step of 5 min at 72 °C. The PCR products are purified on gel using the NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel) according to the manufacturer's protocol.

b. Gateway cloning reactions

To clone *HESO1*, C. De Almeida performed the BP reaction overnight at 25 °C with 100 ng of purified PCR product, 100 ng of donor vector p207 and 1 µL of the BP Clonase® II reaction mix (Invitrogen) in a final volume of 5 µL. The enzymatic reaction is stopped by adding 0.2 µg/µL of proteinase K and incubating the reaction 10 min at 37 °C. The reaction is then used to transform 50 µL of competent *E. coli* cells.

To produce the different tagged versions of HESO1, the entry plasmid generated by C. De Almeida is purified and further used for the LR reaction. 100 ng entry plasmid, 100 ng destination vector and 1 µL of the LR Clonase® II reaction mix (Invitrogen) in a final volume of 5 µL are incubated for 1 hour at 25 °C. The LR reaction is stopped by adding 0.2 µg/µL of proteinase K and incubating the reaction 10 min at 37 °C. This reaction mix is used to transform 50 µL of competent *E. coli* cells.

To produce a C-terminal GFP-tagged version of URT1, H. Scheer carried out the LR reaction following the same protocol, using a pDONR™Zeo containing URT1 5' UTR and the introns (Ferrier, 2013).

c. Escherichia coli cell transformation

50 µL of chemically competent *E. coli* cells are added to the BP/LR reaction mix and incubated for 30 min on ice. The heat shock is performed by incubating the mix for 30 s at 42 °C. The mix is incubated on ice for 1 min and 1 mL of liquid LB medium is added. The cells are then incubated for 1 hour at 37 °C. 100 µL of the suspension are put on LB agar plates supplemented with antibiotics for the selection of the Gateway® entry or destination plasmid. The plasmid is next purified with the NucleoSpin Plasmid QuickPure kit (Macherey-Nagel) according to the manufacturer's instructions.

2. Generation of the TUTases overexpressing lines

a. Agrobacterium tumefaciens cell transformation

50 µL chemically competent *A. tumefaciens* cells (GV3101 strain) are incubated with 100 ng purified plasmid for 30 min on ice. The heat shock is performed at 37 °C for 5 min. The mix is incubated 2 min on ice and 1 mL liquid LB medium is added. The cells are incubated 2 hours at 28 °C. 100 µL of the suspension are put on LB agar plates supplemented with antibiotics for the selection of the Gateway® entry or destination plasmid. The plasmid is next purified with the NucleoSpin Plasmid QuickPure kit (Macherey-Nagel) according to the manufacturer's instructions.

b. Agrobacterium-mediated plant transformation

Arabidopsis thaliana plants at flowering stage (6-7 weeks) are transformed following a protocol previously described (Bernhardt et al., 2012). *A. tumefaciens* transformed with the binary vector containing the insert of interest are grown in a pre-culture of 5 mL LB medium and antibiotics (diluted at 1/10 000) for 16-24 h at 28 °C. A 100 mL culture of LB medium and antibiotics (diluted at 1/10 000) is inoculated with 100 µL of the pre-culture and incubated for 16-24 h at 28 °C.

Agroinfiltration medium is freshly prepared as following: MS medium M0222 at 2.2 g/L; sucrose at 5% w/v; Silvet L-77 at 0.005% v/v; acetosyringone at 200 µM.

A. tumefaciens cells are pelleted by centrifugation for 10 min at 5500 g and resuspended in agroinfiltration medium to obtain an OD of 0.8. The cells are incubated for 1 h in the dark. The plants are dipped into this solution for 30 s. After the floral dip, the plants are placed in the dark for 24 h. The seeds of these plants (T0 generation) are harvested to start screening for positive transformants (T1 generation).

c. Screening of transformants

To screen the T1 generation for transformants with the URT1-GFP, HESO1-GFP or HESO1-RFP insert (pUBC vectors, glufosinate resistance), seeds are sown on soil in trays. After 10-12 days, the seedlings are sprayed twice with glufosinate diluted in water at 0.002% (v/v). The resistant plants are transferred in individual pots at day 14-16 and the expression of the construction is checked by Western blot on flowers.

To screen the T1 generation for transformants with the HESO1-4xmyc insert (pGWB17 vector, hygromycin resistance), seeds are sown on selective agar medium with hygromycin. The plates are placed in the dark at 4 °C for 48 h to allow seed stratification. The plates are then transferred in light

conditions and transformants are transferred on soil in individual pots at day 12-15. The expression of the construction is checked by Western blot on flowers.

To generate stable homozygous lines, the positive T1 transformants are further selected on selective agar medium and the expression of the construct is checked by Western blot on flowers. Rounds of selection are performed until we obtain ~100% resistant plants on selective medium.

3. Western blot analysis (selection of TUTases overexpressing plants)

a. Protein extractions

Flower or leaves are flash-frozen in liquid nitrogen. Proteins are extracted either by SDS-urea protocol or by using TRI-Reagent (Molecular Research Center) followed by acetone precipitation.

For the SDS-urea protocol, tissues are grinded in tubes with glass beads (4 mm diameter, Carl Roth) for 8 s using the mixing device Silamat® S6 (Ivoclar Vivadent). 100 µL SDS-urea-DTT buffer (Tris pH 6.8, 62.5 mM; urea 4 M; SDS 2% (w/v); glycerol 10% (w/v); DTT 100 mM) are added and the tissues are further grinded for 20 s with the Silamat mixing device. The tubes are heated for 10 min at 70 °C and spined 5 min at 16 000 g. The samples can be loaded on an SDS-PAGE gel.

For TRI-Reagent extraction protocol, tissues are grinded in tubes with beads for 8 s with the Silamat mixing device. 300 µL TRI-Reagent are added and the tissues are grinded for 15 s using Silamat. 60 µL chloroform are added and the tubes are vortexed for 10 s. After 15 min incubation at room temperature (RT), the tubes are spined for 15 min at 18 000 g, RT. The aqueous phase containing RNA can be kept. 100 µL ice cold ethanol (100%) are added, the tubes vortexed for 10 s and incubated for 15 min at RT. DNA is precipitated by centrifugating the tubes for 15 min at 18 000 g, RT. The supernatant is transferred into a new tube and proteins are precipitated by adding 3 volumes of acetone (100%). After 5 min incubation on ice, the tubes are spined for 5 min at 5000 g, RT. The supernatant is removed and the protein pellet is washed twice with acetone 80%. 100 µL SDS-urea-DTT buffer (same as above) and glass beads are added into each tube. The tubes are mixed for 15 s with the Silamat mixing device and directly heated for 5 min at 95 °C. After a short spin, the supernatant is transferred into a new tube and Laemmli SDS buffer is added (stock 4x: 250 mM Tris-HCl pH 6.8, 8% SDS, 40% glycerol, 8% ß-mercaptoethanol, 0.02% bromophenol blue).

b. SDS-PAGE

The protein extracts are separated by SDS-PAGE (polyacrylamide gel electrophoresis containing sodium lauryl sulfate). It is composed of a 10% running gel (375 mM Tris-HCl pH 8.8; 0.1% SDS (w/v); 10% polyacrylamide/N,N' methylene bisacrylamide 37.5/1 (v/v)) and of a 4% stacking gel (125 mM Tris-HCl pH 6.8; 0.1% SDS (w/v); 4% polyacrylamide/N,N' methylene bisacrylamide 37.5/1 (v/v)). The gel polymerization is induced by addition of ammonium persulfate (APS) and tetramethylethylenediamine (TEMED). 10-20 µL protein extracts are loaded and the gel is run in migration buffer (25 mM Tris-HCl pH 8.5, 250 mM glycine, 0.1% SDS (w/v)) for 40-70 min at 25 mA / gel.

c. Western blot

Proteins are transferred on a 0.45 µm Immobilon-P PVDF membrane (Millipore), activated for a few seconds in 100% ethanol, by immersion in a transfer buffer (48 mM Tris; 39 mM glycine; 15% ethanol (v/v)) at 4 °C for 90 min at 250 mA (using the Bio-Rad Trans-Blot SD system). After transfer, the membrane is blocked in TBS-T buffer (20 mM Tris-HCl pH 7.4; 150 mM NaCl; 0.2% Tween-20 (v/v)) with 5% (w/v) milk powder for 30 min at RT. After saturation, the membrane is incubated with the primary antibody diluted in TBS-T buffer with 2% (w/v) milk powder (see table below). The unbound antibodies are removed by two washes of 10 min using the TBS-T buffer. The membrane is then incubated with the secondary antibodies diluted in TBS-T buffer with 5% (w/v) milk powder (see table below) for 30 min at RT. Finally, the membrane is washed by three washes of 5 min before revelation. The proteins bound by the antibody are detected by chemiluminescence using Lumi-Light Wester Blotting Substrate (Roche) and luminescence is detected with Fusion FX camera system.

Antibody	Dilution	Incubation time	Secondary antibody
@myc	1/ 10 000	Overnight, 4 °C	@GAM (goat anti mouse)
@URT1	1/10 000	90 min., RT	@GAR (goat anti rabbit)
@GFP	1/50 000	90 min., RT	@GAR (goat anti rabbit)
@GFP-HRP	1/10 000	60-90 min., RT	-

Summary of the antibodies used for Western blots

d. Coomassie staining of PVDF membranes

To detect the proteins transferred on the membrane and to control the loading, PVDF membranes are rinsed with water and incubated overnight with Coomassie staining solution (3 mM Coomassie Brilliant Blue R-250, 45% ethanol; 9% acetic acid). The membrane is then incubated in a destaining solution (45% ethanol, 9% acetic acid). The Coomassie blue dye binds to proteins via hydrophobic bonds.

4. Viral inoculation

PVX and TRV were inoculated from cell culture of infectious clones. GFLV (GT isolate), TMV, PCV, BBWV-1, CMV and TYMV were mechanically inoculated from purified virions, viral RNAs or infectious plasmid. TCV, TuMV, SoMV, AlMV, TBSV, BNYVV, ArMV (*C. quinoa*), TRSV, SLRSV (*C. quinoa*), GFLV (GHu and B844 isolates on *C. quinoa*), CNDV, CLRV, CPMV and RpRSV were inoculated by SAP from infected tissues. TuYV was inoculated on *B. napus* by aphid transmission (Leiser et al., 1992). Grapevine rootstock Kober 5BB were inoculated by heterologous grafting with GFLV- (isolates GHu and B844) or ArMV-infected *C. quinoa* plants. Infected vines were conserved in a greenhouse. GLRaV-1, GLRaV-2, GRSPaV, GPGV, GRGV, GRVFV, GVB, SLRSV, and TBRV originate from vines in infected vineyards. Inoculation conditions and harvesting time are recapitulated in Supplementary Data 1c available from:

https://data.mendeley.com/datasets/k3njd5yfj3/draft?a=bd2e49d5-ac42-4b84-befb-23e6df7e2011

5. RNA extraction

RNA from infected leaves in *A. thaliana* with TYMV, TuMV, GFLV GT or TCV, *C. quinoa* with SoMV, *B. napus* with TuYV, *N. benthamiana* with AlMV, PVX, AlMV, TCV, TBSV, TMV, TRV, PCV, *N. clevelandii* with CMV and *S. oleracea* with BNYVV is extracted as following. Tissues are grinded in liquid nitrogen and ~200 µL are put in tube with glass beads (Carl Roth). 800 µL TRI-Reagent (Molecular Research Center) are added and the tubes are mixed 15 s using the mixing device Silamat® S6 (Ivoclar Vivadent). The cell debris are removed by centrifugation for 5 min at 12 000 g, 4 °C. The supernatant is transferred into a new tube, 160 µL chloroform are added and the tubes are vortexed. After a 10 min incubation at RT, the tubes are centrifuged 10 min at 18 000 g, 4 °C. The upper translucid aqueous phase containing the RNA is transferred into a new tube. The organic phase containing DNA and proteins can be kept at – 20 °C for further extraction. 0.5 volume of isopropanol is added and after vortexing, the tubes are incubated for 10 min at RT. The RNA is precipitated by centrifugation for 10 min at 18 000 g, 4 °C. The RNA pellet is washed with ice cold 75% ethanol and resuspended in 150 µL MilliQ water. Next, 150 µL phenol-chloroform-isoamyl alcohol (25-24-1) are added and after vigorous vortexing, the tubes are centrifuged for 10 min at 18 000 g, 4 °C. The upper translucid aqueous phase containing the RNA is transferred into a new tube. 0.5 µL Glycogen™ (20 mg/µL, Invitrogen), 1/10 volume sodium acetate (3 M, pH 5.2) and 2.5 volumes of ice cold 100% ethanol are added and the samples are put at – 80 °C for at least one hour. The RNA is then precipitated by a 30 min centrifugation step at 18 000 g, 4°C. The RNA pellet is washed twice with ice cold 75% ethanol and resuspended in 21 µL MilliQ water. The RNA is quantified by spectrophotometry (Thermoscientific Nanodrop 2000). The quality is the RNA extracted is next checked by loading 200-400 ng total RNA on a 1% agarose gel.

For some samples, the protocol has been optimized to minimize RNA degradation. For total RNA from BNYVV-infected *S. oleracea*, PCV-infected *N. benthamiana* and SoMV-infected *C. quinoa*, the upper translucid phase obtained after the phenol-chloroform step was further purified using the NucleoSpin® RNA Cleanup kit (Macherey-Nagel), following the manufacturer's instructions, and RNA pellets were resuspended in 15 µL MilliQ water.

Shahinez Garcia (SVQV team, INRAE Colmar) extracted the RNA from infected leaves in *C. quinoa* with ArMV, CLRV, CPMV, RpRSV, GFLV Ghu, GFLV B844, TRSV, SLRSV, *N. benthamiana* with CNDV, *N. benthamiana* B2 with BBWV-1 using the RNeasy plant Mini Kit (Qiagen). For grapevine tissues, the RLC buffer mixed with DTT 25 mM and PVP40 1% (w/v) is used. For herbaceous tissues, the RLT buffer mixed with DTT 25 mM is used. The rest of protocol is carried out following the manufacturer's instructions. The RNA is quantified by spectrophotometry (Thermoscientific Nanodrop 2000). The quality is the RNA extracted is next checked by loading 200-400 ng total RNA on a 1% agarose gel.

In addition to the analysis of total RNAs, viral RNAs from GFLV purified virions (isolates K30 and B844) were analyzed. To extract RNA from purified virions, our colleagues from the INRAE performed a phenol-chloroform purification step, followed by RNA precipitation.

6. Viral accumulation analysis

a. GFP signal quantification

TuMV-GFP infection in *A. thaliana* was visualized by the green fluorescent protein (GFP) signal under UV light (λ = 365 nm). The plants were photographed under UV light. The GFP signal was quantified using ImageJ software (v. 2.0.0-rc-43/1.52n) and the macro "Excessive Green Base" provided by J. Mutterer. For replicates 3 and 4 of TuMV-infected plants that overexpress TUTases, the plants were also photographed under white light. The size of the rosette was quantified using the same macro. To obtain the amount of fluorescence by plant taking into account the size of the rosette, the ratio between the green signal from the GFP and from the rosette was calculated.

b. Quantitative PCR

To quantify TuMV RNA accumulation, 5 µg total RNA from infected A. thaliana was DNase-treated with 5U DNase I (Thermofisher) and 1x DNase I buffer (Thermofisher) in a final volume of 50 µL. The reaction is incubated 30 min at 37°C in a water-bath. The DNase treatment is stopped by adding 1 volume phenol-chloroform-isoamyl alcohol (25-24-1). The DNase-treated RNAs are then further precipitated and purified as above. The quality of the RNA is checked by loading 200-400 ng DNase-treated RNA on a 1% agarose gel. To synthetize complementary cDNA, 1 µg DNase-treated RNA is then mixed with 50 pmol oligo$(dT)_{18}$ (Thermofisher), 50 ng random primers (Thermofisher), 10 nmol dNTP in a final volume of 13 µL. The RNA is denatured for 5 min at 65°C. 200 U SuperScript™ IV Reverse Transcriptase (Invitrogen), 1X SuperScript™ IV Reverse Transcriptase buffer (Invitrogen), 40 U of RNase OUT™ (Invitrogen) and 5 mM DTT are added and the reaction is placed in a thermocycler to complete the reverse transcription as following: 10 min at 50 °C, 10 min at 80 °C.

qPCR is then performed either on TuMV or GFLV RNAs as well as on the control gene *ACT2* of A. thaliana (*AT3G18780*) (primers see 1. Material 6.b.). Technical triplicates are done for each primer couple. 0.1 µL cDNA is mixed with 2.5 µM of each primer and SYBR green 1x (Applied Biosystems) in a final volume of 10 µL. The reaction is then incubated in a LightCycler® 480 II (Roche) and carried out as following: 5 min denaturation at 95°C, followed by 45 cycles of 10 s denaturation at 95°C, 15 s annealing at 58°C, 15 s elongation at 72°C, a final denaturation step of 5 s at 95°C and a final elongation step of 1 min at 55°C. The software of the device (version 1.5.1) allows to obtain the values of the cycle threshold (Ct) that are then further analyzed in a spreadsheet. For each technical triplicate, the relative quantity 2^{-Ct} is calculated, for the target (GFLV or TuMV) and the control (*ACT2*). This value is averaged across the triplicates in the control. Next, for each technical triplicate, the 2^{-Ct} ratio of the target to the control mean is calculated. Finally, the mean of the ratios are used to draw plots representing the relative amount of viral RNA in each plant.

7. 3'RACE-seq library preparation

3'RACE-seq is protocol used to analyze the 3' extremities of viral RNAs and is summed up in Fig. 38. An adapter is ligated to the 3' extremity of RNA and it is then used to initiate the synthesis of the complementary DNA (cDNA). The 3' regions of target RNAs are next amplified by nested-PCR on the cDNA. The nested-PCR allows to improve specificity and to add Illumina adapters for sequencing as

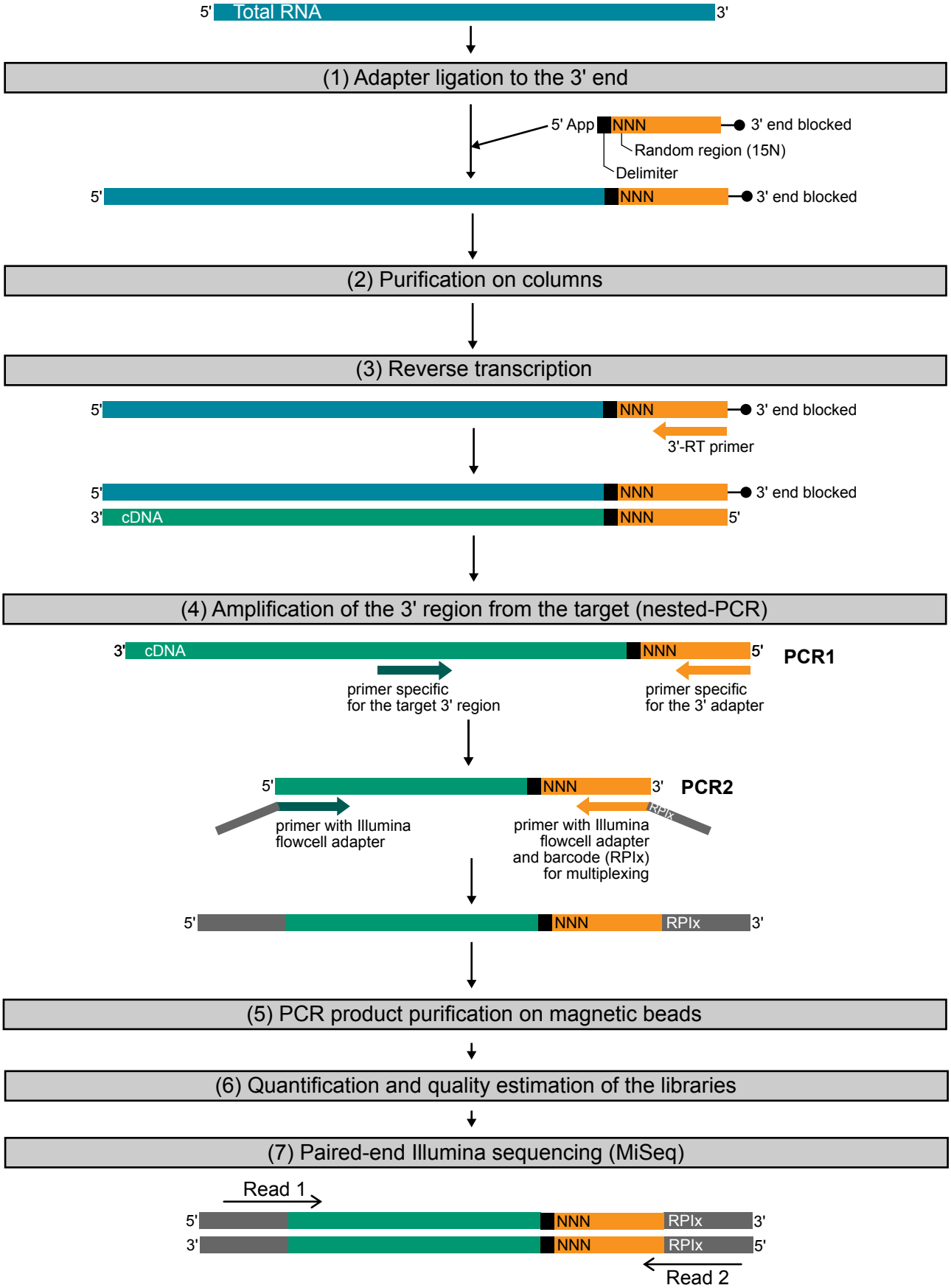


Figure 38: 3'RACE-seq workflow. (1) 3' hydroxy-ends of RNA are ligated to an adapter, which is riboadenylated in 5' and blocked in 3' to prevent self-ligation. **(2)** The ligation products are purified and **(3)** the cDNA synthesis is initiated by a primer complementary to the 3' adapter. **(4)** The 3' regions of target RNAs are amplified by nested-PCR with a forward sequence specific to the RNA and a reverse primer that anneals to the 3' adapter. PCR2 primers also contain Illumina flowcell adapters and barcodes, which allows a multiplexed analysis. **(5)** PCR products are purified on magnetic beads. **(6)** The libraries are quantified and their size is estimated on a Bioanalyzer. **(7)** The libraries are paired-end sequenced on a MiSeq device: the read 1 allows to identify the target and the read 2 provides informations on the templated and non-templated nucleotides at the 3' extremity.

well as barcodes for multiplexing. The libraries are paired-end sequenced on a MiSeq device (Illumina): the read 1 allows to identify the target and the read 2 provides information on the sequence of its 3' extremity.

a. 3' adapter ligation and reverse transcription

For polyadenylated viral RNAs (and GLRaV-1 and CNDV RNAs, first thought as polyadenylated), the ligation is carried out with the classical 3' RACE-seq protocol (Scheer et al., 2021). Ten picomoles of 3' adapter riboadenylated at its 5' extremity (Fig. 38, 1. Material 6.a.) is mixed with 5 µg total RNA and MilliQ water and the RNA is denatured for 5 min at 65 °C. 10 U T4 ssRNA Ligase 1 (NEB) and 1X T4 ssRNA Ligase 1 buffer (NEB) are added (50 µL final volume). The reaction is incubated for 1 hour at 37°C. The ligation products are purified with the NucleoSpin® RNA Cleanup kit (Macherey-Nagel). After elution, 0.5 µL GlycoBlue™ (15 mg/uL, Invitrogen), 1/10 volume sodium acetate (3 M, pH 5.2) and 2.5 volumes of ice cold 100% ethanol are added and the samples are put at – 80 °C for at least one hour. The RNA is then precipitated by a 30 min centrifugation step at 18 000 g, 4°C. The RNA pellet is washed twice with ice cold 75% ethanol and resuspended in 11 µL MilliQ water. The RNA is quantified by spectrophotometry (Thermoscientific Nanodrop 2000). The quality is the RNA after ligation is next checked by loading 200-400 ng total RNA on a 1% agarose gel.

For non-polyadenylated viral RNAs (AlMV, CMV, PCV, SoMV, TMV, TBSV, TCV, TRV, TuYV, TYMV), the ligation protocol was optimized to reduce the potential bias due to strong structures at the 3' extremity of these RNAs. 5 µg total RNA in MilliQ water is denatured at 95 °C for 10 s. 10 pmol of the 3' adapter, 10 U T4 ssRNA Ligase 1 (NEB), 1X T4 ssRNA Ligase 1 buffer (NEB) and 6.25% PEG 8000 are added (50 µL final volume). The reaction is incubated for 1 hour at 37°C and the ligation products are purified as mentioned above.

2 µg of purified ligated RNA are taken for reverse transcription. 50 pmol of 3' RT oligo (1. Material 6.a.), 10 nmol dNTP and RNA are mixed with MilliQ water in 13 µL final volume. The reaction is heated for 5 min at 65 °C and cooled 2 min on ice. 200 U SuperScript™ IV Reverse Transcriptase (Invitrogen), 1X SuperScript™ IV Reverse Transcriptase buffer (Invitrogen), 40 U of RNase OUT™ (Invitrogen) and 5 mM DTT are added and the reaction is placed in a thermocycler to complete the reverse transcription as following: 10 min at 50 °C, 10 min at 80 °C.

b. Nested-PCR amplification and purification of 3' RACE-seq libraries

The 3' region of each viral RNA is amplified by nested-PCR from the cDNA. The PCR is performed in a final volume of 20 µL containing 0.5 µM forward and 0.5 µM reverse primers, 0.2 mM dNTP, 1.5 mM MgCl2, 1 U GoTaq® DNA Polymerase (Promega), 1X GoTaq® DNA Polymerase Green Buffer (Promega). For the first PCR, 1-2 µL cDNA is taken. The forward primer is specific to the target and the reverse primer anneals to the 3' adapter. 1 µL of PCR1 is used for the second PCR where the forward primer is specific to the target and the reverse anneals to the adapter. Both PCR2 primers harbor Illumina adapters and the reverse primer also carries a barcode, which allows to multiplex the samples for the sequencing. All the primers used in this part are listed in "1. Material 6.a.". The PCR1 starts with a denaturation step at 94 °C for 30 s, followed by 20 cycles of denaturation at 94 °C for 20 s, annealing at 50-54 °C for 20 s, elongation at 72 °C for 30-40 s and a final elongation step at 72 °C for 30-4 s. The

PCR2 starts with a denaturation step at 94 °C for 1 min, followed by 20 cycles of denaturation at 94 °C for 30 s, annealing at 51-58 °C for 20 s, elongation at 72 °C for 30-40 s and a final elongation step at 72 °C for 30-40 s. The quality of the PCR amplification is checked by loading 3 µL of PCR2 on a 1.2% agarose gel. The PCR2 products are then purified on homemade magnetic beads (Malek Alioua, IBMP) with a ratio beads:PCR product of 0.8:1 or 1:1 depending on the size of the amplicon. Beads and PCR products are mixed at the desired ratio and incubated for 5 min at RT. The tubes are then placed on a magnetic holder for 5 min. The supernatant is taken away and the beads are washed twice with ice cold 75% ethanol. The beads are left to dry for 1-2 min. Then, the beads are resuspended in 12-15 µL MilliQ water and incubated for 5 min at RT. The tubes are placed back on the magnetic holder for 5 min. Finally, the supernatant containing the retrieved DNA is transferred into a new tube and DNA is quantified by spectrophotometry (Thermoscientific Nanodrop 2000). The amplicon sizes are then checked on a 1.2% agarose gel.

To generate the 3' RACE amplicons of ArMV, CLRV, CPMV, RpRSV, GFLV GHu, GFLV B844, TRSV, SLRSV, CNDV and BBWV, the same protocol as above was used by S. Garcia except for the three following points. The reverse transcription is performed using 200 U SuperScript™ III Reverse Transcriptase (Invitrogen) and 40 U RNAsin® (Promega) as RNase inhibitor per reaction. 30 cycles are done for each PCR. The amplicons are purified on AMPure XP beads (Beckman Coulter), with a ratio beads:PCR product of 0.8:1 or 1:1 depending on the size of the amplicon, according to the manufacturer's instructions.

c. Libraries sequencing

The obtained libraries are quantified using the Qubit fluorometer (Invitrogen™) and their quality is evaluated using the Bioanalyzer 2100 system (Agilent). The samples are pooled at an equimolar ratio. 9 or 10 pM of the final library are used for sequencing. Spike-in sequences with known poly(A) tail lengths are added to assess the quality of the poly(A) tail size estimation during the analysis. 30% of a Phix control v3 library is sequenced in parallel to add complexity to the library. The libraries are paired-end sequenced on a MiSeq™ sequencer (Illumina), with 41 and 111 cycles for read 1 and read 2, respectively.

8. 5'RACE-seq library preparation

To determine the sequence of GFLV minus strand 5' region, we set up a "5'RACE-seq" strategy which combines cDNA synthesis using the 5' RACE kit (Invitrogen™) and our 3'RACE-seq method (Article Fig. S6). Target specific cDNA are synthesized using the kit according to manufacturer's instructions. The advantage of this kit is that the obtained cDNA strand is C-tailed at its 3' extremity. We take advantage of this for the nested-PCR. The reverse primer for the first PCR is a custom primer derived from our 3' RACE-seq adapter (see 1. Material 6.a.). This custom primer contains a GI sequence that anneals to the C-tail at the 3' extremity of the cDNA strand. Like our 3'RACE-seq adapter, it also contains a delimiter sequence, to distinguish between the adapter sequence and the 3' end of the RNA (the C-tail will be removed during the analysis, see Chapter "5' and 3' RACE-seq data analysis"), an 8 random nucleotide sequence, to allow the detection of PCR duplicates and a sequence that will permit the annealing the PCR1 primer (identical to the sequence found on the 3' adapter).

a. cDNA synthesis

5 µg total RNA are used to synthetize a target specific cDNA strand following the manufacturer's instructions of the kit 5' RACE System for Rapid Amplification of cDNA Ends, version 2.0 (Invitrogen™). The target specific primers are listed in "1. Material 6.a".

b. Nested-PCR amplification and purification of 5' RACE-seq libraries

Three rounds of PCR were performed to amplify the 5' region of GFLV RNAs. The PCR is performed in a final volume of 20 µL containing 0.5 µM forward and 0.5 µM reverse primers, 0.2 mM dNTP, 1.5 mM MgCl2, 1 U GoTaq® DNA Polymerase (Promega), 1X GoTaq® DNA Polymerase Green Buffer (Promega). For the first PCR, 1 µL cDNA is taken. The forward primer is specific to the target and the reverse primer anneals to the 3' C-tail of the cDNA. 1 µL of PCR1 and PCR2 are used for PCR2 and PCR3, respectively. In both PCR2 and PCR3, the forward primer is specific to the target and the reverse anneals to the adapter. In the PCR3, the primers harbor Illumina adapters and the reverse primer also carries a barcode which allows to multiplex the samples for the sequencing. All the primers used in this part are listed in "1. Material 6.a.". Of note, the reverse primers used for PCR3 also are the TruSeq Illumina primers. The PCR1 starts with a denaturation step at 94 °C for 1 min, followed by 25 cycles of denaturation at 94 °C for 30 s, annealing at 55-65 °C for 30 s, elongation at 72 °C for 40 s and a final elongation step at 72 °C for 40 s. The PCR2 starts with a denaturation step at 94 °C for 1 min, followed by 20 cycles of denaturation at 94 °C for 30 s, annealing at 55-65 °C for 30 s, elongation at 72 °C for 40 s and a final elongation step at 72 °C for 40 s. The PCR3 starts with a denaturation step at 94 °C for 1 min, followed by 5 cycles of denaturation at 94 °C for 30 s, annealing at 55-65 °C for 30 s, elongation at 72 °C for 40 s and a final elongation step at 72 °C for 40 s. The quality of the PCR amplification is checked by loading 3 µL of PCR2 on a 1.2% agarose gel. The PCR2 products are then purified on homemade magnetic beads (Malek Alioua, IBMP) with a ratio beads:PCR product of 0.6:1 or 0.8:1 depending on the size of the amplicon. The purification on beads is performed as for 3'RACE-seq libraries. DNA is quantified by spectrophotometry (Thermoscientific Nanodrop 2000). The amplicon sizes are then checked on a 1.2% agarose gel.

c. Libraries sequencing

The 5'RACE-seq libraries are further prepared and sequenced like the 3'RACE-seq libraries.

9. 5' and 3'RACE-seq data analysis

Base callings attributed by the MiSeq Control Software v 2.5 (Illumina) and extracted for further analysis. The analysis is conducted by a set of homemade scripts using python (v2.7), biopython (v1.63) and regex (v2.4) libraries. First, reads with low quality of base calling (=< Q10 in 3'RACE-seq and =< Q20 in 5'RACE-seq) are removed. Then, sequences with identical nucleotides in the random sequence (15 in 3'RACE-seq and 8 in 5'RACE-seq) are deduplicated. The 20 first nucleotides of the read 1 are issued to identify the corresponding target, with one mismatch tolerated. Identified read 1 and their corresponding read 2 are extracted. Reads 2 that contain the delimiter sequence are selected and the delimiter sequence, as well as the random sequence, are trimmed.

Then, four options are possible, depending on the type of sequencing (5' or 3'RACE-seq) and on the target: "A", "B1B2", "B2" and "5'RACE". In 3'RACE-seq, the option "A" is used for non-polyadenylated RNAs. The read 2 is mapped onto the reference sequence and the length and composition of non-templated nucleotides are analyzed and classified according to nine categories: no tail, onlyU, onlyA, onlyC, onlyG, U-rich, A-rich, C-rich, G-rich. For the onlyX classes, the tail is strictly composed of the corresponding nucleotide whereas for the X-rich classes, the tail contains at least 70% of the corresponding nucleotide, therefore the detected tail is at least three nucleotides long.

For 3'RACE-seq on polyadenylated RNAs, the poly(A) tail length prevents the read 2 mapping. First, as we only sequence 111 nucleotides (including 20 nucleotides of the adapter), we cannot map read 2 if the tail is long. Moreover, the low quality of base calling after the poly(A) hinders the efficient mapping of the read 2. If the reference sequence is available, we can still map the read 2 but it is often a small proportion of the reads restricted to RNAs with no or short poly(A) tails. The pipeline contains two options ("B1B2" or "B2") depending on whether or not we have the reference sequence of the target. In option "B1B2", the read 2 is mapped onto the reference sequence and the length and composition of non-templated nucleotides are analyzed and classified according to the following categories: polyA, polyAandU, polyAandG, polyAandC, polyU, polyG, polyC, polyAandhetero and no tail. The four first classes regroup poly(A) tails with or without other non-templated nucleotides at the 3' end of the poly(A) tail. The polyU, polyG and polyC characterize non-templated tails found after non-polyadenylated RNA. Finally, the polyAandhetero class contains tails with a poly(A) followed by heteropolymeric tails. For the remaining reads that cannot be mapped the pathway "B2" continues the analysis: poly(A) tails containing at least 8 adenosines are searched and the length and composition of the non-templated tail is analyzed as above. Results of the B1 and B2 analyses are combined into a final file. If the reference sequence is not available, the analysis is only performed with the option B2, described above.

Finally, the 5'RACE option is used to analyzed the 5'RACE-seq data on GFLV RNAs. The analysis follows the same steps as the B1B2 option of polyadenylated RNAs, instead, a poly(U) tail is searched. The classes of tails obtained in the end are: polyU, polyUandA, polyUandG, polyUandC, polyA, polyG, polyC and polyUandhetero.

10. Plots and statistical analyses

All plots were generated using R (v4.0.2) and ggplot2 R (v3.3.5) on RStudio (v1.3.1093). Boxplots show the median and the 1st and 3rd quartiles (Q1 and Q3). The minimal and maximal values are shown by the whiskers, calculated as following: min = Q1 – 1.5x(Q3-Q1) and max = Q3 + 1.5x(Q3-Q1). The values that are lower or higher than these limits are represented by individual points.

For all statistical analyses, a p-value of 0.05 is defined as the threshold of significance. Statistical analyses are conducted using the pairwise Wilcoxon rank-sum test (two-tailed). Detailed results of the statistical analyses are provided in Supplemental table 5.

Résumé en français

1. Introduction

1. Introduction générale sur les phytovirus

Dans la nature, les plantes doivent faire face à des conditions climatiques défavorables et à des attaques de pathogènes, tels que des virus, des champignons ou des bactéries. Les virus sont notamment responsables d'importants dommages sur les cultures au niveau mondial et représentent la moitié des maladies émergentes des cultures, ce qui est similaire à la part des maladies virales humaines. De plus, le changement climatique contribue à l'émergence des maladies virales dans le secteur agricole car il modifie notamment la distribution et l'abondance des arthropodes, qui sont les principaux vecteurs viraux dans la nature (Anderson et al., 2004 ; Hogenhout et al., 2008).

90% des virus infectant des plantes ont un génome à ARN, et parmi eux, la moitié a un génome à ARN simple brin de polarité positive (ARN sb(+)) (ICTV, 2021), qui est directement reconnu par la machinerie de traduction de la cellule végétale. Les extrémités 5' et 3' du génome viral peuvent posséder des caractéristiques similaires aux ARN messagers (ARNm) cellulaires, telle que la coiffe en 5' et la queue poly(A) en 3'. D'autres caractéristiques ont été décrites comme la protéine virale liée au génome (*genome-linked viral protein*, VPg) en 5', ou des structures ARN en 3' telles que des structures similaires aux ARN de transfert (*tRNA-like structure*, TLS) ou des structures en tige boucle. La figure 1 récapitule les différentes caractéristiques des extrémités 5' et 3' publiées pour les ARN viraux.

De nombreux ARN viraux sont polyadénylés et comme pour les ARNm, la queue poly(A) des ARN viraux peut assurer une fonction de protection et participe donc au maintien de l'intégrité du génome viral. La queue poly(A) peut aussi être impliquée dans la formation de structures qui inhibent la dégradation 3'-5' (Torabi et al., 2021b). Pour certains virus, elle est également impliquée dans la réplication du génome et la régulation de la traduction (Geng et al., 2019 ; Paul and Wimmer, 2015). Pour les ARN viraux non polyadénylés, la 3' *untranslated region* (3' UTR) peut être caractérisée par la formation de structures ARN tel que la TLS, comme c'est le cas pour l'ARN du *turnip yellow mosaic virus*. Les TLS ont été beaucoup étudiées pour leur proximité structurale avec les ARNt et afin de mieux comprendre leur rôle dans le métabolisme viral. Il a été montré que les TLS sont impliquées dans la maintenance de l'intégrité du génome viral, la réplication et la traduction (Dreher, 2010). En outre, d'autres ARN viraux présentent des structures ARN stables en 3' dont l'organisation tridimensionnelle a, pour certaines, été résolue, tel que pour l'ARN du *turnip crinkle virus* (Fig. 6). Ces structures en 3' sont notamment importantes pour le maintien de l'intégrité du génome et la traduction (Simon, 2015).

2. La dégradation des ARN, un mécanisme de défense antiviral

Pour se défendre contre les infections virales, les plantes disposent de stratégies telles qu'une réponse immunitaire innée, le mécanisme de *silencing*, ou encore la répression traductionnelle (Wu et al., 2019). Les voies de dégradation des ARN participent également à la défense antivirale en ciblant

les ARN viraux. De manière générale, la dégradation des ARN est cruciale pour réguler l'expression des gènes et le maintien de l'homéostasie cellulaire dans tous les organismes. Certains des facteurs de dégradation des ARNm étant impliqués également dans la dégradation des ARN viraux, je vous résume ci-après les principales étapes de dégradation des ARNm dans le cytosol des eucaryotes.

Principales voies de dégradation des ARNm chez les eucaryotes

Dans le cytosol des eucaryotes, le raccourcissement de la queue poly(A), appelée déadenylation, constitue la première étape des voies de dégradation d'un ARNm et permet sa dégradation de 5' en 3' ou de 3' en 5' (Fig. 9). Après la déadenylation, la coiffe présente en 5' est clivée par le complexe d'élimination de la coiffe (Fig. 9), rendant ainsi accessible l'extrémité 5'-phosphate de l'ARN aux exoribonucléases 5'-3' telles que Xrn1 chez la levure et l'humain ou XRN4 chez les plantes (Fig. 9).

Alternativement, les ARN peuvent être dégradés par la voie de 3' en 5' qui implique notamment le complexe de l'exosome (Fig. 9). Ce complexe multiprotéique est constitué d'un cœur central, Exo9, de neuf sous-unités qui interagit avec des facteurs variés selon la localisation intracellulaire. L'activité catalytique de l'exosome chez les levures et les métazoaires est assurées par des exoribonucléases associées à Exo9, comme Dis3 dans le cytosol ou Rrp6 dans le noyau (Chlebowski et al., 2011). Il est à noter que chez les plantes, Exo9 s'associe également avec des exoribonucléases mais possède une activité polyribonucléotide phosphorylase (Sikorska et al., 2017). En plus des exoribonucléases Dis3 ou Rrp6, les cofacteurs d'Exo9 incluent des hélicases ARN, des protéines de liaison à l'ARN, des poly(A) polymérases et des protéines d'échafaudage. Dans le cytosol, le complexe Ski est un partenaire majeur de l'exosome. Le complexe Ski est composé de l'hélicase Ski2, la protéine Ski3 contenant des répétitions tétratricopeptides et deux copies de la protéine Ski8 à motifs WD40 (Brown et al., 2000 ; Halbach et al., 2013 ; Kowalinski et al., 2016).

En plus du complexe de l'exosome, la dégradation de 3' en 5' peut être assurée par l'exoribonucléase DIS3L2 (Fig. 9). Cette enzyme, initialement identifiée chez Arabidopsis où elle est appelée SUPPRESSOR OF VARICOSE (SOV) (Zhang et al., 2010), a été largement étudiée chez *S. pombe* et les mammifères où elle cible des ARN non-codants (Chang et al., 2013 ; Haas et al., 2016 ; Łabno et al., 2016 ; Nowak et al., 2017 ; Pirouz et al., 2016, 2019 ; Ustianenko et al., 2013, 2016) et les ARNm (Lubas et al., 2013 ; Malecki et al., 2013).

Interaction entre les virus et les voies de dégradation des ARN de l'hôte

En ce qui concerne les interactions entre les facteurs de dégradation des ARN et les ARN viraux chez les plantes, il existe relativement peu d'études focalisées sur les mécanismes moléculaires sous-jacents. Une étude récente a par exemple montré que les protéines LSM1 et PAT1, impliquées dans l'élimination de la coiffe des ARNm, sont requises pour la multiplication du *turnip mosaic virus* (Zuo et al., 2022). Les études sur l'exoribonucléase XRN4 ont décrit des conséquences parfois contradictoires de l'absence de cette enzyme sur le métabolisme des ARN viraux. Certaines études ont montré le rôle antiviral de XRN4 (Cheng et al., 2007 ; Jiang et al., 2018 ; Lee et al., 2015 ; Li and Wang 2018 ; Ma et al., 2019 ; Peng et al., 2011). Cependant, d'autres études ont rapporté que l'absence de XRN4 est néfaste pour la réplication virale, suggérant un rôle proviral de cette enzyme (Vogel et al., 2011 ; Gy et

al., 2007 ; Lee et al., 2015). Les mécanismes de dégradation des ARN viraux chez les plantes sont donc encore peu connus et restent à être caractérisés.

3. L'uridylation : un mécanisme clé de régulation du métabolisme des ARN

Acteurs et mécanismes de l'uridylation

La dégradation des ARN chez les eucaryotes peut impliquer des modifications post-transcriptionnelles telles que la polyadénylation des ARN non-codants dans le noyau. L'ajout d'uridines à l'extrémité 3' des ARN, appelée uridylation, est également fréquemment détectée. Cette modification post-transcriptionnelle est conservée chez tous les eucaryotes testés à l'exception de *S. cerevisiae* et émerge comme un mécanisme clé de régulation du métabolisme des ARN. L'uridylation cible aussi bien les ARN non-codants que les ARNm. La conséquence principale de l'uridylation des ARN est leur dégradation, mais elle est également impliquée dans la maturation de certains ARN non-codants (De Almeida et al., 2018b ; Liudkovska and Dziembowski, 2021 ; Warkocki et al., 2018a ; Yu and Kim, 2020 ; Zigáčková and Vaňáčová, 2018).

L'addition d'uridines est catalysée par des enzymes appelées uridylyltransférases terminales (TUTases), qui sont constituées d'un domaine catalytique, comprenant un domaine nucléotidyltransférase et un domaine poly(A) polymérase, et parfois de régions additionnelles impliquées dans la régulation de leur activité (Fig. 10) (De Almeida et al., 2018b ; Zigáčková and Vaňáčová, 2018). Chez Arabidopsis, deux TUTases ont été caractérisées à ce jour : HEN1 SUPPRESSOR 1 (HESO1), qui uridyle principalement les petits ARN et les fragments issus du clivage d'ARNm par le complexe RISC, et UTP:RNA URIDYLYL TRANSFERASE 1 (URT1), qui est la principale responsable de l'uridylation des ARNm (Ren et al., 2012, 2014 ; Sement et al., 2013).

L'uridylation des ARN viraux

A l'instar des ARN cellulaires, l'uridylation a aussi été détectée sur des ARN viraux. En effet, une étude a détecté de l'uridylation 3' terminale sur des ARN viraux provenant de virus infectant des champignons, des animaux et des plantes (Huo et al., 2016). Plus récemment, il a été mis en évidence que la TUTase CDE-1 de *C. elegans* agit comme un facteur antiviral contre le virus Orsay (Le Pen et al., 2018). Cette étude montre également que les TUTases TUT4 et TUT7 des mammifères uridylent les ARNm du virus Influenza A, entrainant leur dégradation. Les auteurs de cette étude proposent donc l'uridylation des ARN viraux comme nouveau mécanisme antiviral. Pourtant, les données mécanistiques relatives aux processus moléculaires sous-jacents de l'uridylation des ARN viraux et ses conséquences restent rares et, à ce jour, il n'existe aucune étude portant sur l'uridylation et les mécanismes associés à l'uridylation des ARN de phytovirus.

Données préliminaires sur l'uridylation des ARN viraux chez les plantes

Avant le début de ma thèse, mon équipe disposait de données préliminaires relatives à l'uridylation des ARN de six virus à ARN sb(+) infectant des plantes. En utilisant la stratégie de 3'RACE-seq (3' *Rapid Amplification of cDNA Ends-sequencing*), ils ont analysé les profils d'uridylation de ces ARN viraux. Le 3'RACE-seq permet d'étudier les ajouts de nucléotides en 3' d'ARN cibles en combinant l'amplification par PCR et le séquençage haut-débit (Scheer et al., 2020). Notre équipe utilise cette

méthode pour l'étude de cibles variées telles que les ARNm et les fragments d'ARNm issus du clivage par le complexe RISC (Scheer et al., 2021, Zuber et al., 2018). Leurs données de 3'RACE-seq sur les ARN viraux montrent que tous les ARN sont uridylés avec des pourcentages variant de 0,5% pour l'ARN de l'ORMV à plus de 80% pour les deux ARN du GFLV (Fig. 12A). De plus, ces données révèlent deux profils d'uridylation distincts : les ARN du GFLV sont mono-uridylés tandis que les autres ARN analysés ont des extensions d'uridines entre 1 et 30 Us, avec une majorité d'extensions entre 1 et 3 Us (Fig. 12B). Il convient de noter que le profil d'uridylation particulier détecté pour les ARN du GFLV n'a, à notre connaissance, jamais été observé pour un ARN chez les plantes. Ces résultats préliminaires sont particulièrement intéressants puisqu'ils suggèrent une diversité dans les mécanismes moléculaires associés à l'uridylation des ARN viraux chez les plantes. De manière similaire à ce qui a été montré pour les ARN eucaryotes, l'uridylation des ARN viraux pourrait impliquer différents facteurs et avoir des conséquences variées.

2. Objectifs de thèse

C'est dans ce contexte qu'a été initié mon projet de thèse dont le but est de contribuer à une meilleure compréhension des mécanismes moléculaires liés à l'uridylation des ARN viraux chez les plantes. Au vu des connaissances actuelles portant sur l'uridylation des ARN viraux chez les plantes, des questions clés sont soulevées :

- Quelle est la diversité des profils d'uridylation présents sur les ARN viraux chez les plantes ?
- Quelles sont les enzymes responsables de l'uridylation des ARN viraux, sont-elles des facteurs de l'hôte ou des facteurs viraux ?
- Quel est l'impact de l'uridylation sur la stabilité des ARN viraux ?
- Quel est le rôle de l'uridylation des ARN dans les interactions hôte-virus ?

Le premier objectif de mon travail de doctorat était d'étudier la diversité des profils d'uridylation des ARN viraux chez plantes. Notre équipe a initié une collaboration avec le groupe d'Olivier Lemaire à l'INRAE de Colmar (France) qui est spécialisé dans l'étude des virus infectant la vigne. Au total, nous avons analysé par 3'RACE-seq les profils d'uridylation de plus de trente ARN de phytovirus, qui sont des représentants des principales familles de phytovirus à ARN sb(+), avec un focus sur des espèces appartenant à la famille des *Secoviridae*, dont fait partie le GFLV.

Le deuxième objectif de ma thèse était d'identifier les facteurs impliqués dans l'uridylation des ARN viraux. Pour cela, je me suis concentrée sur trois virus modèles, deux virus avec des ARN polyadénylés, GFLV et TuMV, un virus avec un ARN non polyadénylé, TCV. Ces trois virus infectent et se répliquent dans notre plante modèle *Arabidopsis thaliana*. J'ai évalué l'impact de l'inactivation de HESO1 et URT1, les deux TUTases identifiées à ce jour chez Arabidopsis, sur les profils d'uridylation des ARN viraux de ces virus.

Enfin, le dernier objectif de mon travail était d'initier l'étude de l'impact de l'uridylation sur le métabolisme viral. J'ai d'abord évalué l'impact d'une absence ou d'une surexpression des TUTases d'Arabidopsis sur l'accumulation virale *in planta*. J'ai généré des plantes transgéniques surexprimant

HESO1 alors que des lignées surexprimant URT1 étaient déjà disponibles dans l'équipe. En parallèle, j'ai commencé l'analyse de l'accumulation et de l'uridylation des ARN viraux dans des plantes mutées pour les facteurs impliqués dans la dégradation des ARN dans le cytosol.

Les résultats que j'ai collectés au cours de ma thèse constituent une ressource sur l'uridylation des ARN viraux chez les plantes, nécessaire pour mieux appréhender les rôles pro- et antiviraux potentiels de l'uridylation dans le métabolisme viral. Mes données fournissent également un aperçu des acteurs impliqués dans l'uridylation des ARN viraux chez les plantes. Mon travail de thèse ouvre la voie à une meilleure compréhension du lien potentiel entre uridylation et dégradation des ARN viraux chez les plantes.

3. Résultats et Discussion

1. Étude globale de l'uridylation 3' terminale des ARN de phytovirus : identification de variations extrêmes et de profils d'uridylation spécifiques à certains ARN viraux

L'uridylation 3' terminale sur des ARN viraux chez les champignons, les métazoaires et les plantes a été observée par séquençage Sanger, pour seulement quelques clones (Huo et al., 2016). Afin d'analyser les profils d'uridylation des ARN viraux chez les plantes par séquençage haut-débit, nous avons utilisé la stratégie de 3'RACE-seq qui permet de déterminer les nucléotides de la région 3' d'un ARN cible. Cette technique consiste en l'amplification par PCR de la région 3' terminale de l'ARN cible qui est ensuite séquencée par séquençage haut-débit.

Nous avons choisi des virus représentant les principales familles de phytovirus à ARN sb(+). Vingt-neuf virus couvrant onze des quinze familles de phytovirus à ARN sb(+) ont été choisis, représentés sur la figure 13. Nos collaborateurs pour ce projet étant spécialisés dans les virus infectant la vigne, plus de la moitié des virus que nous avons analysés infectent cette plante.

Limites de la stratégie de 3'RACE-seq

La stratégie de 3'RACE-seq (Fig. 38) a plusieurs limitations qu'il est important de prendre en considération pour interpréter correctement les résultats. Tout d'abord, le 3'RACE-seq repose sur la ligation d'un adaptateur en 3' des ARN se terminant par un groupement hydroxyl. Ainsi, la réaction de ligation peut être empêchée si l'ARN ne se termine pas par un groupe hydroxyl, notamment en cas d'une modification 3' terminale. De plus, de fortes structures ARN peuvent gêner la ligase. Pour pallier ce problème et sur les conseils du Dr. Anthony Gobert, j'ai optimisé le protocole de ligation pour les ARN non polyadénylés en dénaturant à plus haute température. Enfin, il est important de noter que notre stratégie ne permet pas de distinguer les différents types d'ARN viraux qui partagent les mêmes extrémités 3' (génomique et subgénomique notamment).

Les profils d'uridylation des ARN viraux sont variés

Nos données révèlent que les profils d'uridylation des ARN de phytovirus sont très variés (Article fig. 2), allant de 0,2 à 90% pour les ARN de GFLV. Nous avons détecté une stricte mono-uridylation sur les queues poly(A) des ARN du GFLV qui est indépendante des TUTases de l'hôte et constitue un nouveau type d'extrémité jusqu'ici inconnue sur des ARN viraux (Article fig. 2, 7 et 9). De plus, cette mono-uridylation des queues poly(A) n'est pas répandue pour les ARN d'autres espèces virales de la famille des *Secoviridae*, puisqu'elle est restreinte aux ARN du GFLV et de l'ArMV, une espèce virale très proche phylogénétiquement (Article fig. 8).

Certains des ARN viraux étudiés qui se terminent par une structure semblable à un ARNt cellulaire (TLS) peuvent être aminoacylés (Dreher, 2010), prévenant donc la ligation et la détection par 3'RACE-seq. La population d'ARN viraux se terminant par un TLS détectée correspond donc à des ARN non-aminoacylés. Nos données montrent que les ARN se terminant par un TLS non-aminoacylés sont peu uridylés (Article fig. 2), ce qui reflète une faible accessibilité aux TUTases et une potentielle compétition avec d'autres enzymes de modification 3' terminales, telles que les aminoacyl-ARNt synthétases responsables de l'aminoacylation des ARNt dans la cellule.

De plus, certains ARN non polyadénylés ont montré des taux d'uridylation relativement hauts (environ 10%). Dans ce cas, nous avons détecté de l'uridylation sur des ARN de pleine longueur mais également des ARN tronqués (Article fig. 3, 4 et S2). Ces profils d'uridylation sont conservés dans les trois plantes analysées et sont donc probablement des signatures d'attaques par des endoribonucléases ou des exoribonucléases 3'-5' générant ces intermédiaires de dégradation qui sont ensuite uridylés. Dans le cas des ARN polyadénylés, nous avons observé que l'uridylation a lieu sur des ARN aux queues poly(A) inférieures à 30 adénosines (Article fig. 5), comme dans le cas d'ARNm ayant subi une étape de déadénylation (Scheer et al., 2021 ; Zuber et al., 2016). Cela suggère que les ARN viraux polyadénylés sont déadénylés avant d'être uridylés. Dans les deux cas, ARN polyadénylé et non-polyadénylé, nos résultats tendent à indiquer que l'uridylation marque des intermédiaires de dégradation.

Autres modifications 3' terminales des ARN viraux

Nous avons aussi pu observer d'autres nucléotides que des uridines ajoutés à l'extrémité 3' des ARN viraux. Nous avons détecté de la mono-adénylation et de la mono-cytidylation sur des ARN décrits comme ayant une structure TLS en 3', ainsi que des queues riches en A sur l'ARN du GLRaV-1 (Fig. 16, 17 et 18). Ces observations témoignent de la diversité des ajouts de nucléotides en 3' des ARN viraux et potentiellement d'implications moléculaires variées dans le métabolisme des ARN viraux.

De plus, le 3'RACE-seq permet de mesurer la taille des queues poly(A) par séquençage haut-débit. Pour les vingt-huit ARN viraux polyadénylés que nous avons analysés, nous avons tracé la distribution de la taille des queues poly(A). Nos résultats montrent que la distribution de la taille des queues poly(A) est très diverse (Fig. 19). Certains ARN viraux ont une queue poly(A) qui présente un pic entre 20 et 30 adénosines. Dans ce cas, il est possible que la distribution de la queue poly(A) soit influencée par la liaison de protéines telles que les *poly(A) binding proteins* (PABP). En effet, l'empreinte d'une PABP sur la queue poly(A) chez la levure est d'environ 28 adénosines (Webster et al., 2018). La taille de la queue poly(A) des ARN viraux pourrait également varier au cours du cycle viral et les

différentes populations de queues poly(A) d'un ARN donné pourraient correspondre à des populations d'ARN viraux à différentes étapes du cycle viral. Il a par exemple été monté chez *C. elegans* et en cellules mammifères que les ARNm avec des queues poly(A) courtes sont traduits plus efficacement (Lima et al., 2017). De manière similaire, les queues poly(A) courtes des ARN viraux pourrait correspondre à des ARN en cours de traduction. En accord avec ces données, une étude *in vitro* a montré que les queues poly(A) inférieures à 30 adénosines favorisent la traduction de l'ARN du *wheat yellow mosaic virus* (Geng et al., 2019). Globalement, nos données constituent une ressource utile pour étudier les queues poly(A) des ARN viraux chez les plantes, qui pourrait être intéressante pour des travaux ultérieurs visant à étudier la synthèse et le rôle de cette caractéristique 3' terminale pour les ARN viraux chez les plantes.

2. HESO1 et URT1 ont des fonctions distinctes sur les ARN viraux

HESO1 et URT1 ont des activités différentes

Nous avons identifié l'implication de HESO1 et URT1, les TUTases identifiées chez Arabidopsis dans l'uridylation de l'ARN non polyadénylé du TCV. Dans nos analyses, nous distinguons l'ajout d'uridines entre des extensions contenant uniquement des uridines ou des extensions contenant majoritairement des uridines mais aussi d'autres nucléotides. HESO1 et URT1 ont toutes deux montré une activité dans l'ajout des extensions ne contenant que des uridines, tandis que HESO1 est principalement responsable des extensions composées majoritairement d'uridines (Article fig. 4A).

En outre, nos analyses sur l'uridylation de l'ARN du TCV dans des plantes mutées pour les TUTases HESO1 et URT1 ont montré que ces deux enzymes ont des préférences distinctes pour le nucléotide 3' terminal. URT1 semble préférer une adénosine 3' terminale tandis que HESO1 uridyle des substrats se terminant par d'autres nucléotides qu'une adénosine (Article fig. 4). Cette préférence a déjà été décrite dans le cas des micro ARN (Tu et al., 2015).

HESO1 et URT1 sont également impliquées dans l'uridylation de l'ARN polyadénylé du TuMV. Les deux TUTases peuvent ajouter des uridines mais contrairement à l'ARN du TCV, URT1 est majoritairement responsable de l'uridylation de cet ARN (Article fig. 6A). En outre, nous avons montré que HESO1 et URT1 uridylent des populations de queues poly(A) différentes. HESO1 uridyle des queues poly(A) dont la taille est inférieure à 10 adénosines tandis que URT1 cible des queues poly(A) mesurant moins de 20 adénosines (Article fig. 6B et C).

Nos données ont aussi montré que la longueur des extensions ajoutées par HESO1 et URT1 sur l'ARN du TuMV varie. En effet, URT1 ajoute des extensions de 1-2 uridines alors que les extensions ajoutées par HESO1 sont plus longues, suggérant une processivité plus importante (Fig. 21). Plusieurs articles ont décrit cette processivité distincte entre ces deux enzymes sur des micro ARN et des fragments issus du clivage par RISC (Tu et al., 2015 ; Zhao et al., 2012 ; Zuber et al., 2018).

Nos analyses des activités de HESO1 et URT1 sur les ARN viraux ont montré que ces TUTases ont des préférences de substrat et des processivités différentes, en accord avec les données précédemment publiées sur leur activité sur d'autres substrats. Ces résultats pourraient donc suggérer des fonctions distinctes de HESO1 et URT1 dans le métabolisme des ARN viraux chez les plantes.

HESO1 impacte négativement l'accumulation virale

Une étude chez *C. elegans* et en cellules de mammifères a révélé que l'uridylation affecte négativement le niveau d'accumulation virale (Le Pen et al., 2018). A ce jour, il n'existe à ma connaissance aucune étude menée chez les plantes. Pour étudier cela, j'ai évalué l'impact de l'absence ou la surexpression de HESO1 ou URT1 sur l'accumulation du TuMV.

J'ai montré que l'absence des TUTases n'impacte pas l'accumulation du TuMV (Fig. 23). Bien que je n'aie pas détecté d'effet de l'absence des TUTases d'Arabidopsis sur l'accumulation virale, il n'est pas exclu que la surexpression de ces enzymes puisse impacter l'accumulation virale. J'ai donc également analysé l'accumulation virale dans des plantes surexprimant HESO1 ou URT1 dans deux fonds génétiques différents : $SOV^{inactif}$ and SOV^{actif}.

SOV est l'homologue de DIS3L2, une exoribonucléase 3'-5' dont l'activité est stimulée par l'uridylation 3' terminale (Łabno et al., 2016 ; Lubas et al., 2013 ; Reimão-Pinto et al., 2016 ; Thomas et al., 2015 ; Ustianenko et al., 2013, 2016). Columbia (Col-0) est l'accession d'Arabidopsis la plus employée pour les études expérimentales et a été utilisée pour toutes les expériences de ce travail. Dans la majorité des accessions d'Arabidopsis, le résidu 705 de SOV est une proline alors que c'est une arginine chez Col-0 (Zhang et al., 2010). Les auteurs de cette étude ont montré que cette mutation abolit, ou du moins gêne sa fonction dans la dégradation des ARN. De ce fait, si une voie de dégradation des ARN uridylés par SOV existe chez les plantes, il est probable qu'elle ne soit pas active ou optimale dans l'accession Col-0, qui peut être considérée comme un mutant SOV. Pour évaluer l'impact d'une version fonctionnelle de SOV dans la dégradation des ARN médiée par l'uridylation, j'ai utilisé des plantes qui surexpriment HESO1 ou URT1 dans un fond génétique exprimant une version active de SOV (SOV^{actif}).

Mes données indiquent que la surexpression de HESO1 impacte négativement l'accumulation du TuMV (Fig. 25 et 27). Il est à noter que cet effet est significatif dans le fond génétique SOV^{actif} et aux stades précoces de l'infection systémique entre 7 et 9 jours post-inoculation. De plus, nous pouvons observer que la progression de l'infection virale est également plus lente dans les plantes surexprimant HESO1 (Fig. 26 et 28). Par contraste, je n'ai pas détecté d'effet reproductible robuste de la surexpression de URT1 sur l'accumulation virale globale (Fig. 25 et 27).

Je dispose également de données préliminaires sur l'accumulation du TCV dans des plantes surexprimant les TUTases. Le suivi des symptômes entre 6 et 13 jours post-inoculation a indiqué que la surexpression de HESO1 semble retarder l'apparition de symptômes systémiques, surtout dans le fond SOV^{actif} (Tableau 5).

Il a été montré pour les cellules humaines et de *S. pombe* que DIS3L2, participe à la dégradation 3'-5' des ARN uridylés (Chang et al., 2013, Lubas et al., 2013, Malecki et al., 2013, Thomas et al., 2015, Ustianenko et al., 2013, 2016). Par conséquent, l'uridylation des ARN viraux assurée par HESO1 pourrait stimuler la dégradation 3'-5' par SOV.

Globalement, l'analyse de l'accumulation virale a montré que la surexpression de HESO1 a un effet négatif sur l'accumulation virale, au contraire de URT1 pour laquelle mes données présentent plus de variabilité. Ces résultats restent à confirmer par l'utilisation de lignées supplémentaires indépendantes mais sont tout de même prometteurs et pourraient suggérer un rôle antiviral de HESO1.

4. L'uridylation des ARN du GFLV : une nouvelle fonction de l'uridylation ?

Nous avons mis à jour un profil d'uridylation tout à fait unique pour un ARN chez les plantes avec la mono-uridylation de la queue poly(A) des ARN du GFLV (Article fig. 2 et 5).

Nous avons montré que l'uridine en 3' du brin (+) n'est pas ajoutée par les TUTases HESO1 et URT1 (Article Fig. 7). Par ailleurs, cette uridine est également détectée sur les ARN du GFLV qui sont encapsidés (Article fig. 9). Cela pourrait signifier que cette uridine est encodée par le génome viral. Nous avons donc déterminé les séquences des régions 5' et 3' du brin complémentaire, le brin (-), qui est synthétisé lors de la réplication du génome viral. Nous avons détecté une uridine en 3' du brin (-) qui est complémentaire au premier nucléotide du brin (+) et aucun nucléotide non codé ajouté. En 5' du brin (-), nous avons détecté une queue poly(U), ce qui implique d'une partie de la queue poly(A) est codée. Par contre, nos données n'ont pas révélé d'adénosine en 5' du brin (-) qui indiquerait que l'uridine en 3' du brin (+) soit codée également (Article fig. 7B et S5).

Bien que les ARN du GFLV ne soient pas uridylés par les TUTases d'Arabidopsis, il n'est à ce stade pas exclu que ces enzymes puissent avoir un effet indirect sur l'accumulation virale. J'ai donc quantifié l'accumulation virale dans des plantes sauvages et mutées pour URT1 ou HESO1 et le double mutant. A 14 jours post-inoculation, mes données n'indiquent aucun impact de l'absence des TUTases sur l'accumulation du GFLV (Fig. 22).

A ce jour, nous ne savons pas quelle activité est responsable de l'uridylation des ARN du GFLV. Bien qu'il y ait d'autres gènes pouvant potentiellement coder pour des TUTases dans le génome d'Arabidopsis (De Almeida et al., 2018a), notre hypothèse principale reste que l'ARN polymérase ARN dépendante codée par le génome du GFLV ajoute cette uridine 3' terminale. En effet, des activités nucleotidyltransférases terminales ont été décrites pour d'autres polymérases virales (Arnold et al., 1999 ; Behrens et al., 1996 ; Fullerton et al., 2007 ; Neufeld et al., 1994 ; Ranjith-Kumar et al., 2001 ; Rohayem et al., 2006 ; Smallwood and Moyer, 1993 ; Tomar et al., 2006 ; Tvarogová et al., 2019 ; Wang et al., 2013b ; Wu et al., 2014). Il serait donc intéressant de tester l'activité nucleotidyltransférase terminale de l'ARN polymérase du GFLV, protéine 1^{E}.

A ce stade de l'étude, nous ne pouvons affirmer quelle fonction joue l'uridine en 3' des ARN du GFLV. Une des hypothèses est que cette uridine pourrait être ajoutée pendant la réplication de l'ARN viral (Fig. 37). De manière mutuellement non exclusive avec l'hypothèse précédente, la mono-uridylation des ARN du GFLV pourrait également assurer un rôle de protection de l'extrémité 3'. En effet, la queue poly(A) est la cible de déadénylases et il a été montré *in vitro* que la présence d'une uridine est suffisante pour ralentir le processus de déadenylation (Scheer et al., 2021).

La prochaine étape du projet vise à mieux comprendre l'importance biologique de cette uridine. Nos collaborateurs ont donc commencé à produire des transcrits infectieux qui se terminent soit par une queue poly(A) mono-uridylée soit par une queue poly(A) uniquement. L'objectif est de tester l'efficacité d'infection de ces transcrits et de caractériser les potentielles optimisations *in vivo* des extrémités 3' des ARN du GFLV.

5. Etude de la relation entre l'uridylation et la dégradation des ARN viraux dans le cytosol

La fonction principale de l'uridylation chez les eucaryotes est d'induire la dégradation des ARN. Cela a également été proposé pour les ARN viraux chez *C. elegans* et en cellules mammifères (Le Pen et al., 2018). Pour analyser les mécanismes liés à la dégradation des ARN viraux uridylés chez les plantes, j'ai analysé l'accumulation du TuMV et les ajouts de nucléotides en 3' de l'ARN dans des plantes mutées pour l'exoribonucléase 5'-3' XRN4 ou le cofacteur de l'exosome SKI2 (*xrn4-3* et *ski2-6*), des mutants catalytiques de la sous-unité RRP41 de l'exosome (*rrp41*^{TYA} et *rrp41*^{TYA*}) et des plantes exprimant une version active de SOV (Col-0 SOV).

Mes données ont révélé une accumulation d'ARN du TuMV tronqués et uridylés en 3' dans le mutant *ski2-6* (Fig. 33A, B et C). Ce résultat suggère l'implication de l'exosome dans la dégradation des ARN viraux, recruté par le complexe SKI, au niveau des ARN uridylés. De plus, ces extensions sont plus longues que celles observées en 3' des queues poly(A) de l'ARN du TuMV (Fig. 33D). En accord avec les données précédemment mentionnées sur la processivité de HESO1, cette enzyme est un candidat potentiel pour l'ajout de ces extensions. Une étude très récente a justement montré la connexion entre l'action des TUTases d'Arabidopsis, notamment HESO1, et la dégradation 3'-5' d'ARNm dans le cytosol (Wang et al., 2022).

Il convient également de noter que l'accumulation virale globale du TuMV ne change pas dans le mutant *ski2-6*, cela signifie donc que la dégradation de l'ARN viral ne dépend pas uniquement de l'action de SKI2 et du complexe SKI. Ces fragments tronqués uridylés sont très probablement le résultat de la voie 3'-5' de dégradation. L'ARN du TuMV est déadénylé, puis grignoté et uridylé. Les cycles successifs de rognage-uridylation donnent lieu à des fragments de tailles différentes. Ce résultat prometteur suggère un rôle de SKI2 et de l'exosome cytosolique dans la dégradation des ARN viraux uridylés.

4. Conclusion générale

Les résultats obtenus pendant cette thèse ont permis d'établir une ressource sur l'uridylation des ARN de phytovirus, mais également sur les modifications 3' terminales des ARN en général comme la polyadénylation. Nous avons identifié une variabilité importante dans les profils d'uridylation, suggérant que des mécanismes moléculaires variés pourraient être à l'œuvre. En accord avec cette hypothèse, nous avons montré que l'uridylation est assurée par des facteurs différents selon les ARN viraux : alors que HESO1 et URT1, les deux TUTases caractérisées chez Arabidopsis, sont responsables de l'uridylation des ARN du TuMV et de TCV, d'autres facteurs sont impliqués dans l'uridylation des ARN de GFLV. Nous avons montré que la surexpression de HESO1 impacte négativement l'accumulation virale et nous disposons de d'indices moléculaires du lien entre uridylation et dégradation des ARN viraux. Nos résultats ont donc permis de dresser un inventaire de l'uridylation des ARN de phytovirus et d'amorcer la compréhension de mécanismes moléculaires liés à cette modification fréquente des ARN viraux.

Supplementary data

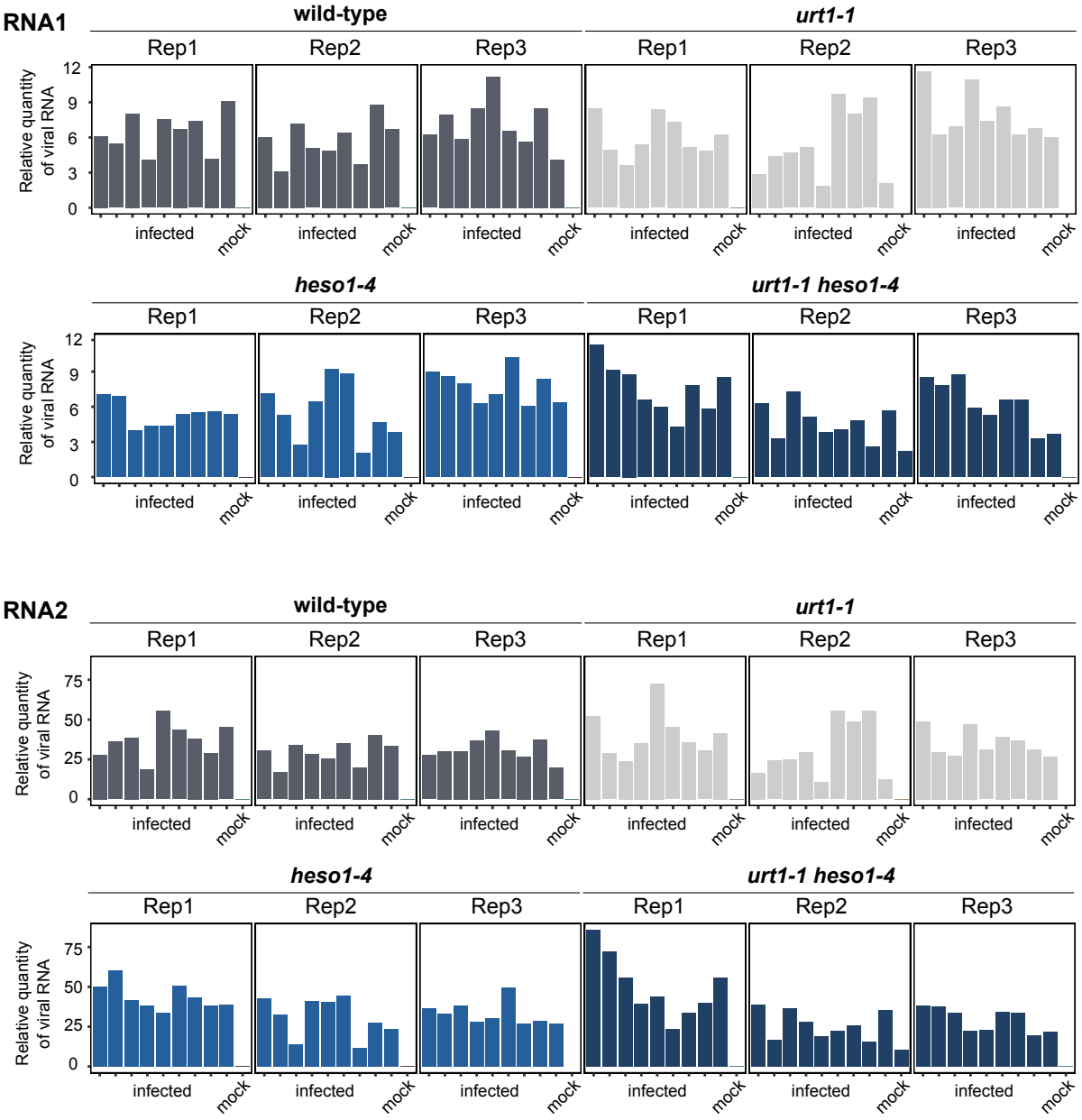


Figure S1 Quantification of GFLV RNA accumulation in wild-type and TUTase mutants by RT-qPCR at 14 dpi. Barplot showing the relative quantity of GFLV RNA1 and RNA2 determined by RT-qPCR in each plant of three biological replicates. The viral RNA quantity has been normalized to Arabidopsis *ACT2* gene. Nine infected and one mock plants per genotype for three independent biological replicates have been quantified.

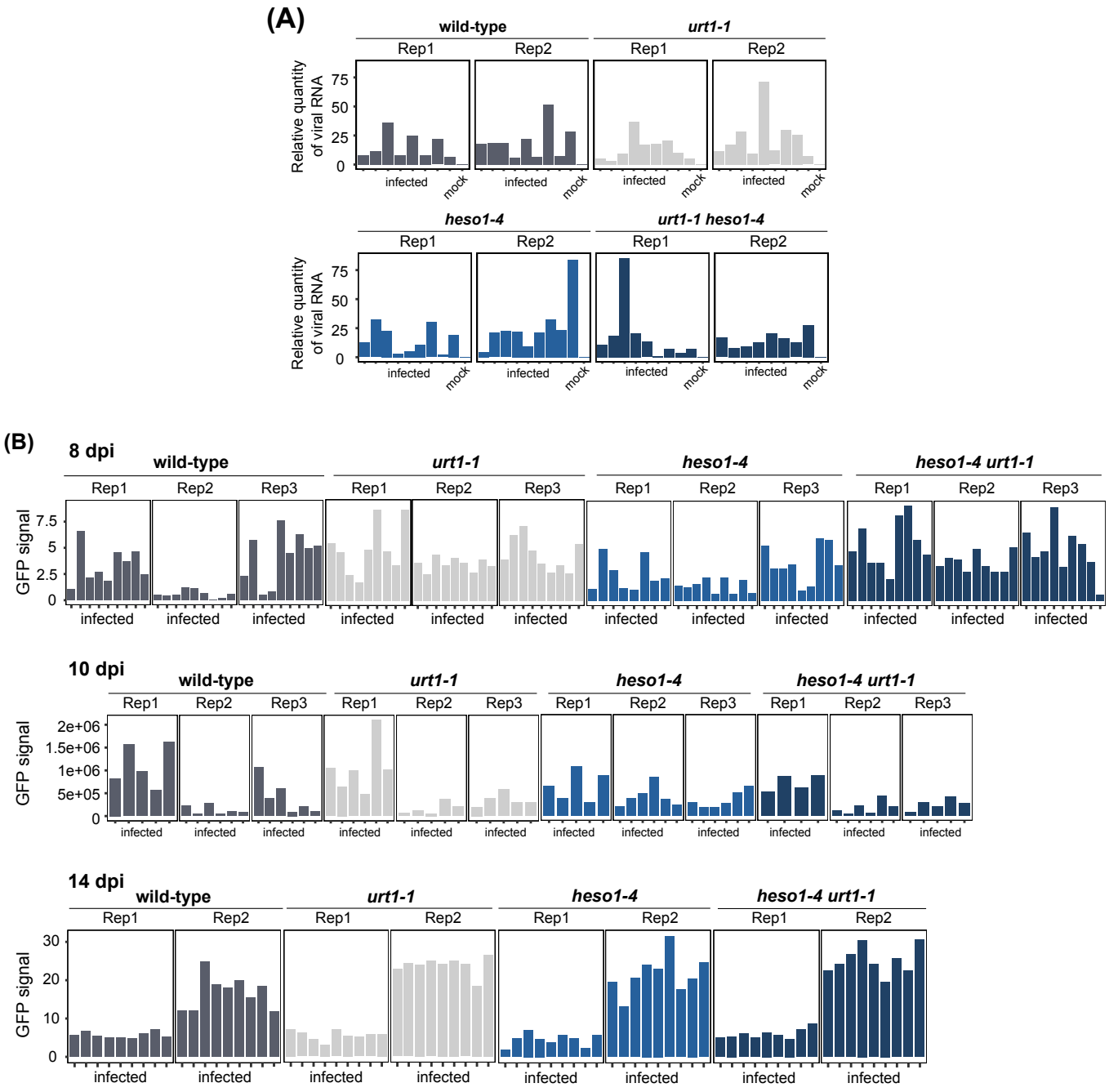


Figure S2 Quantification of TuMV accumulation in wild-type and TUTase mutants. (A) Barplot showing the relative quantity of TuMV RNA determined by RT-qPCR in each plant of two biological replicates at 14 dpi. Nine infected plants per genotype have been quantified. The viral RNA quantity has been determined for nine plants per genotype and normalized to Arabidopsis *ACT2* housekeeping gene. **(B)** Barplot showing the quantification of the GFP signal in each infected plant of two or three independent biological replicates at 8, 10 and 14 dpi. At 8, 10 and 14 dpi, nine, four to six and nine infected plants per genotype have been quantified, respectively. Replicates 1 and 2 at 8 and 14 dpi correspond to the same batches of plants photographed at two different timepoints. By contrast, the three replicates at 10 dpi correspond to independent replicates.

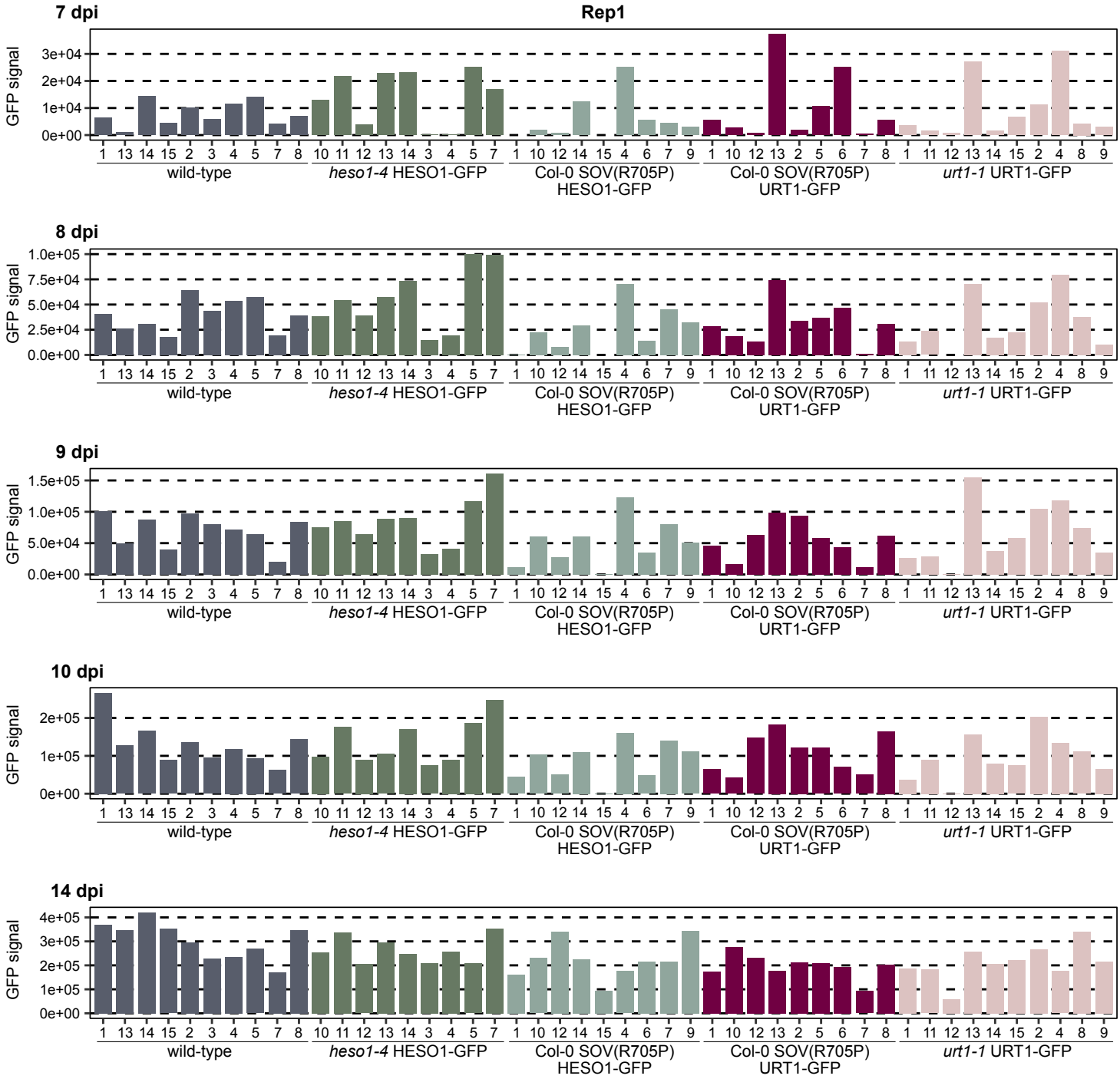


Figure S3 GFP signal quantification in TuMV-GFP infected plants (rep1). Barplot showing the quantification of the GFP signal in each infected plant of replicate 1 at 7, 8, 9, 10 and 14 dpi. 9-10 infected plants per genotype have been quantified.

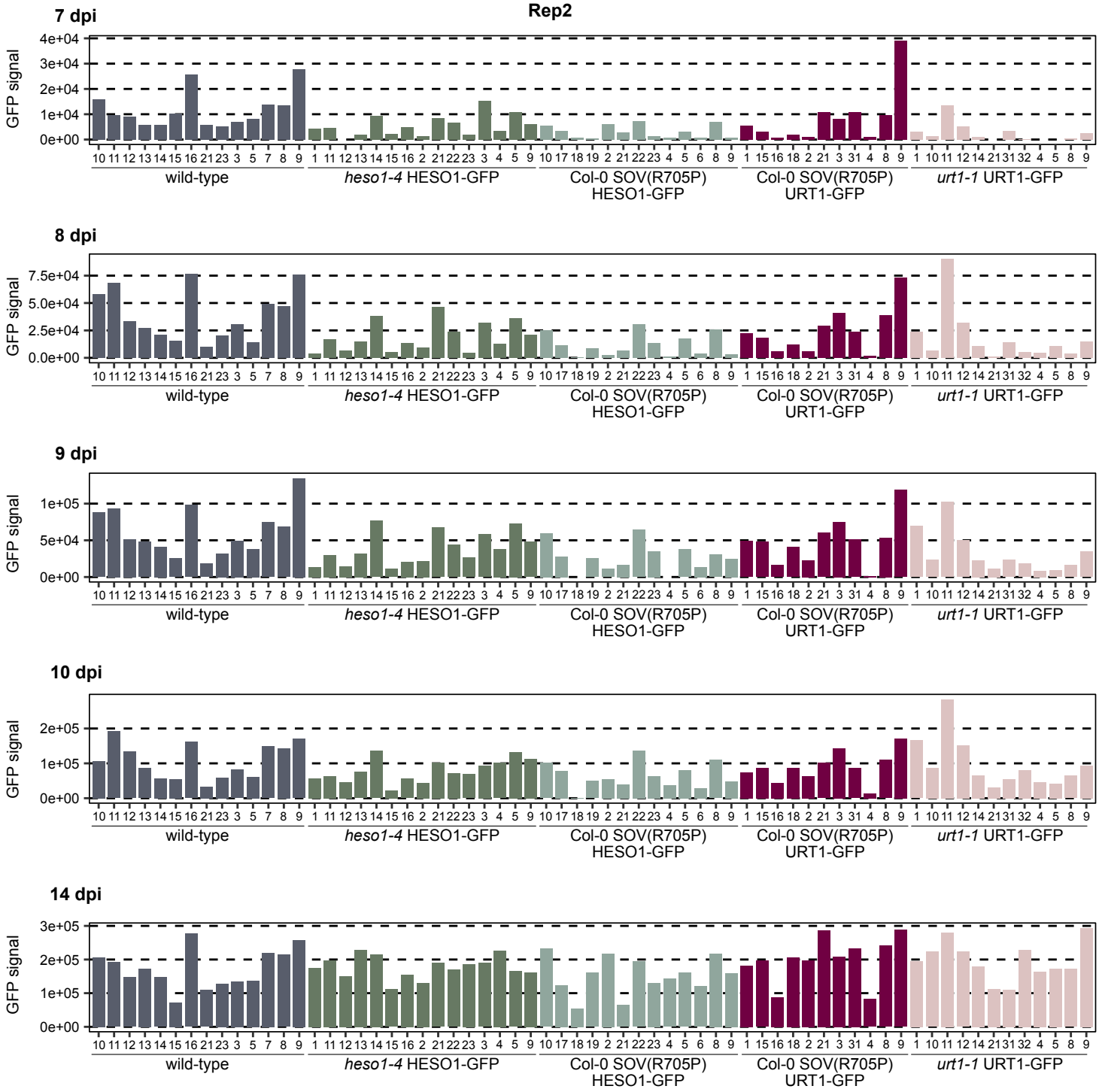


Figure S4 GFP signal quantification in TuMV-GFP infected plants (rep2). Barplot showing the quantification of the GFP signal in each infected plant of replicate 2 at 7, 8, 9, 10 and 14 dpi. 11-15 infected plants per genotype have been quantified.

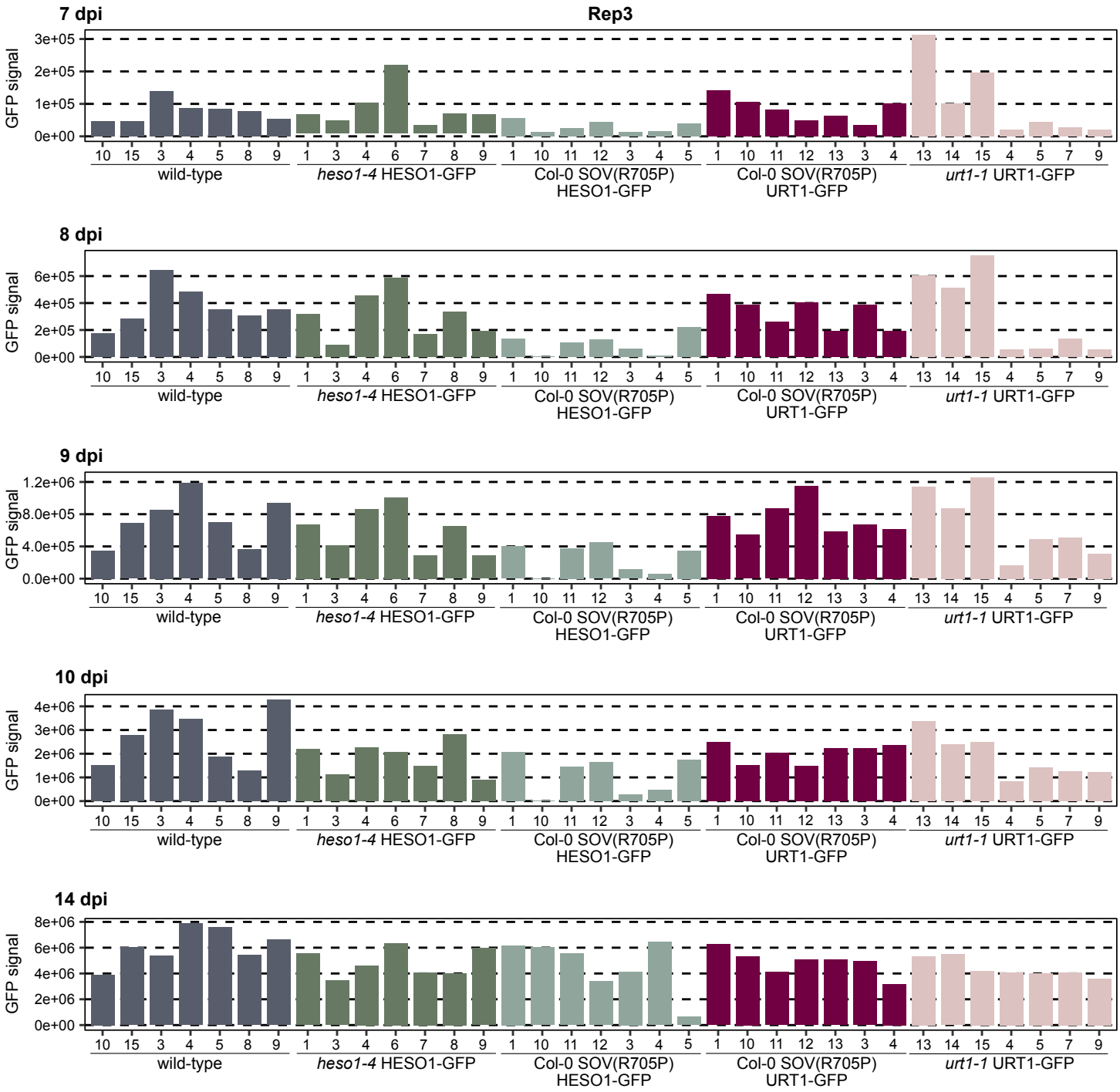


Figure S5 GFP signal quantification in TuMV-GFP infected plants (rep3). Barplot showing the quantification of the GFP signal in each infected plant of replicate 3 at 7, 8, 9, 10 and 14 dpi. 7 infected plants per genotype have been quantified.

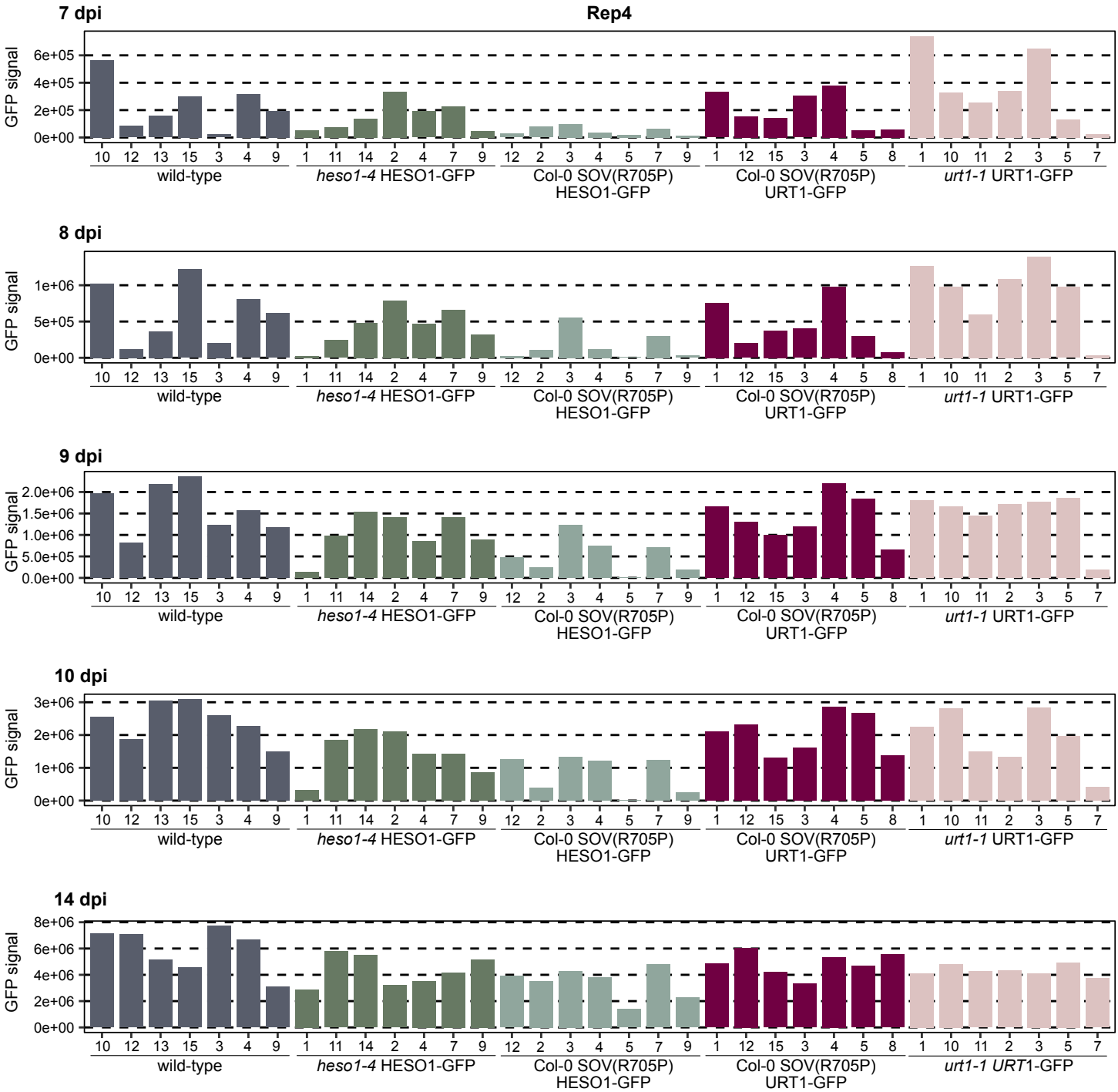


Figure S6 GFP signal quantification in TuMV-GFP infected plants (rep4). Barplot showing the quantification of the GFP signal in each infected plant of replicate 4 at 7, 8, 9, 10 and 14 dpi. 7 infected plants per genotype have been quantified.

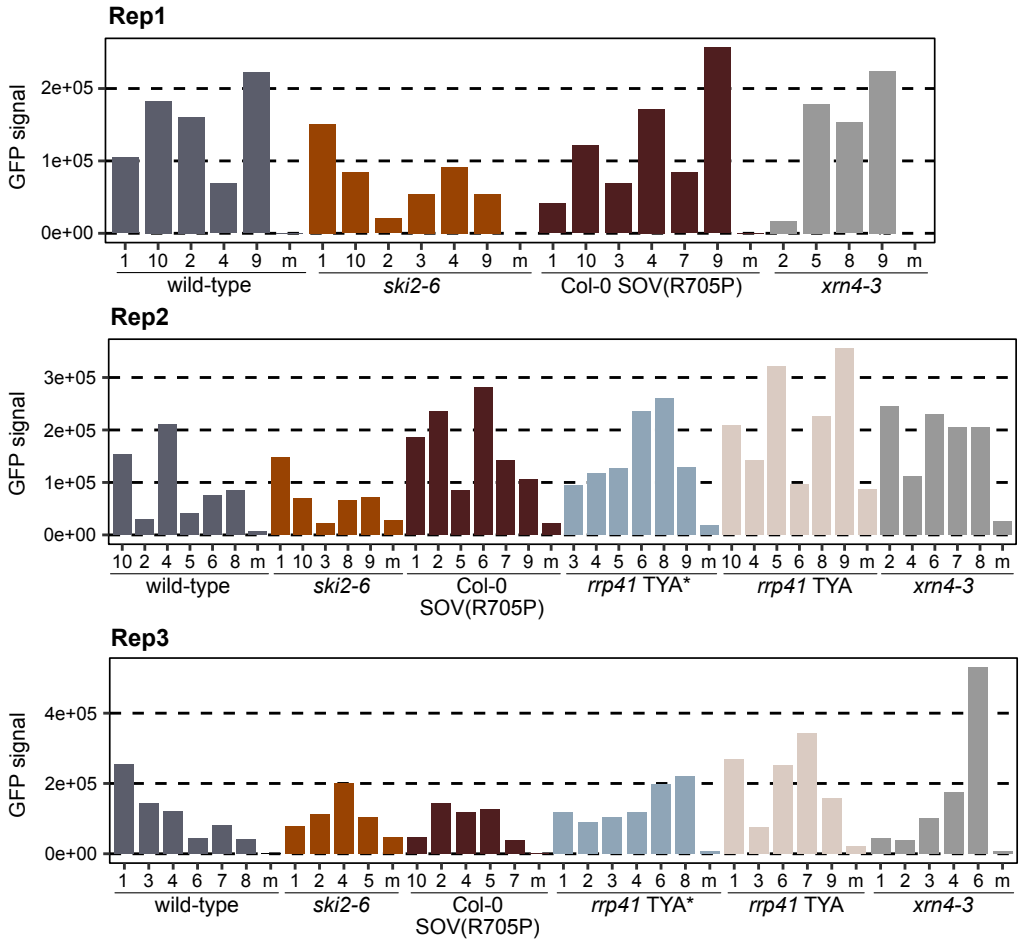


Figure S7 Quantification of TuMV accumulation in wild-type and cytosolic decay mutants. Barplots showing the quantification of the GFP signal in each infected plant by replicate at 10 dpi. 4-6 infected plants per genotype for 2 or 3 independent biological replicates have been quantified. The mock plant is indicated by the letter "m" for each genotype.

Supplementary tables 1-4 related to the article are available from:

https://data.mendeley.com/datasets/k3njd5yfj3/draft?a=bd2e49d5-ac42-4b84-befb-23e6df7e2011

Viral accumulation in urt1-1, heso1-4, heso1-4 urt1-1							
GFLV RNA1 14 dpi				GFLV RNA2 14 dpi			
	Col-0	heso1-4	urt1-1		Col-0	heso1-4	urt1-1
heso1-4	0.99	-	-	heso1-4	0.87	-	-
urt1-1	0.99	0.99	-	urt1-1	0.87	0.87	-
urt1-1heso1-4	0.99	0.99	0.99	urt1-1heso1-4	0.87	0.87	0.87
Col-0	heso1-4	urt1-1	heso1-4urt1-1	Col-0	heso1-4	urt1-1	urt1-1heso1-4
a	a	a	a	a	a	a	a
TuMV 14 dpi qPCR				TuMV 14 dpi GFP			
	Col0	heso1-4	urt1-1		Col-0	heso1-4	urt1-1
heso1-4	0. 94	-	-	heso1-4	0.719	-	-
urt1-1	0. 94	0. 94	-	urt1-1	0.073	0.166	-
urt1-1heso1-4	0. 94	0. 94	0. 94	urt1-1heso1-4	0.043	0.102	0.719
Col0	heso1-4	urt1-1	urt1-1heso1-4	Col0	heso1-4	urt1-1	urt1-1heso1-4
a	a	a	a	a	ab	ab	b
TuMV 10 dpi GFP				TuMV 8 dpi GFP			
	Col0	heso1-4	urt1-1		Col0	heso1-4	urt1-1
heso1-4	0.7	-	-	heso1-4	0.4578	-	-
urt1-1	0.74	0.7	-	urt1-1	0.0057	0.0203	-
urt1-1heso1-4	0.74	0.7	0.7	urt1-1heso1-4	0.0014	0.0119	0.6554
Col0	heso1-4	urt1-1	urt1-1heso1-4	Col0	heso1-4	urt1-1	urt1-1heso1-4
a	a	a	a	a	a	b	b

Viral accumulation in overexpressing lines (IntDen)									
TuMV 7 dpi					TuMV 8 dpi				
	Col-0	HESO1	SOV HESO1	SOV URT1		Col-0	HESO1	SOV HESO1	SOV URT1
HESO1	0.61	-	-	-	HESO1	0.3727	-	-	-
SOV HESO1	0.05	0.22	-	-	SOV HESO1	0.0014	0.0255	-	-
SOV URT1	0.68	0.73	0.2	-	SOV URT1	0.6605	0.8095	0.0239	-
URT1	0.32	0.64	0.61	0.61	URT1	0.3116	0.7808	0.1660	0.7237
Col-0	HESO1	SOV HESO1	SOV URT1	URT1	Col-0	HESO1	SOV HESO1	SOV URT1	URT1
a	a	a	a	a	a	a	b	a	ab
TuMV 9 dpi					TuMV 10 dpi				
	Col-0	HESO1	SOV HESO1	SOV URT1		Col-0	HESO1	SOV HESO1	SOV URT1
HESO1	0.559	-	-	-	HESO1	0.523	-	-	-
SOV HESO1	0.039	0.138	-	-	SOV HESO1	0.067	0.254	-	-
SOV URT1	0.731	0.731	0.098	-	SOV URT1	0.669	0.669	0.089	-
URT1	0.559	0.902	0.386	0.731	URT1	0.523	0.987	0.254	0.669
Col-0	HESO1	SOV HESO1	SOV URT1	URT1	Col-0	HESO1	SOV HESO1	SOV URT1	URT1
a	ab	b	ab	ab	a	a	a	a	a

TuMV 14 dpi				
	Col-0	HESO1	SOV HESO1	SOV URT1
HESO1	0.7	-	-	-
SOV HESO1	0.7	0.7	-	-
SOV URT1	0.7	0.7	0.7	-
URT1	0.7	0.94	0.7	0.7
Col-0	HESO1	SOV HESO1	SOV URT1	URT1
a	a	a	a	a

Viral accumulation in overexpressing lines (Ratio)									
TuMV 7 dpi					TuMV 8 dpi				
	Col-0	HESO1	SOV HESO1	SOV URT1		Col-0	HESO1	SOV HESO1	SOV URT1
HESO 1	0.7006	-	-	-	HESO1	0.9320	-	-	-
SOV HESO 1	0.0026	0.0126	-	-	SOV HESO1	0.0032	0.0083	-	-
SOV URT1	0.4399	0.3964	0.0013	-	SOV URT1	0.9459	0.8342	0.0027	-
URT1	0.4399	0.4399	0.0126	0.6697	URT1	0.3784	0.3784	0.0144	0.3784
Col-0	HESO1	SOV HESO1	SOV URT1	URT1	Col-0	HESO1	SOV HESO1	SOV URT1	URT1
a	a	a	a	a	a	a	b	a	ab
TuMV 9 dpi					TuMV 10 dpi				
	Col-0	HESO1	SOV HESO1	SOV URT1		Col-0	HESO1	SOV HESO1	SOV URT1
HESO 1	0.5342	-	-	-	HESO1	0.0739	-	-	-
SOV HESO 1	0.0026	0.0197	-	-	SOV HESO1	0.0013	0.0685	-	-
SOV URT1	0.5342	0.1533	3.3E-5	-	SOV URT1	0.6347	0.0739	0.006	-
URT1	0.7052	0.4415	0.0026	0.8388	URT1	0.5342	0.5342	0.0685	0.5680
Col-0	HESO1	SOV HESO1	SOV URT1	URT1	Col-0	HESO1	SOV HESO1	SOV URT1	URT1
a	ab	b	ab	ab	a	a	a	a	a

TuMV 14 dpi				
	Col-0	HESO1	SOV HESO1	SOV URT1
HESO1	0.074	-	-	-
SOV HESO1	0.015	0.351	-	-
SOV URT1	0.910	0.62	0.015	-
URT1	0.062	0.910	0.411	0.062
Col-0	HESO1	SOV HESO1	SOV URT1	URT1
a	a	a	a	a

Viral accumulation in decay factors at 10 dpi					
	Col-0	SOV	ski2-6	TYA	TYA*
SOV	0.6792	-	-	-	-
ski2-6	0.3563	0.2078	-	-	-

TYA	0.0341	0.0688	0.0045	-	-
TYA*	0.2760	0.5752	0.0232	0.1720	-
xrn4-3	0.3002	0.4714	0.0688	0.2760	0.7810
Col0	SOV	ski2-6	TYA	TYA*	xrn4-3
ab	abc	a	c	bc	abc

Supplementary table 5: Statistical analyses performed in this manuscript. P-values and groups of significance (two-tailed Wilcoxon statistical test, p-value<0.05) are shown for experiments on viral accumulation in TUTases mutants (Fig. 22, 23), TUTases overexpressing lines (Fig. 25, 27), and cytosolic decay pathways mutants (Fig. 30).

Bibliography

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic Acids Res. *46*, 6257–6270.

Adams, M.., Candresse, T., Hammond, J., Kreuze, J.F., Martelli, G.P., Namba, S., Pearson, M.N., Ryu, K.H., Saldarelli, P., and Yoshikawa, N. (2009). ICTV Virus Taxonomy Profile: Betaflexviridae 2009.

Adams, M.J., Adkins, S., Bragard, C., Gilmer, D., Li, D., MacFarlane, S.A., Wong, S.M., Melcher, U., Ratti, C., and Ryu, K.H. (2017). ICTV Virus Taxonomy Profile: Virgaviridae. J. Gen. Virol. *98*, 1999–2000.

Agranovsky, A.A. (2021). Structure and Expression of Large (+)RNA Genomes of Viruses of Higher Eukaryotes. Biochemistry. (Mosc). *86*, 248–256.

Agranovsky, A.A., Koonin, E. V., Boyko, V.P., Maiss, E., Frötschl, R., Lunina, N.A., and Atabekov, J.G. (1994). Beet Yellows Closterovirus: Complete Genome Structure and Identification of a Leader Papain-like Thiol Protease. Virology *198*, 311–324.

Ahlquist, P., and Kaesberg, P. (1979). Determination of the length distribution of poly(A) at the 3' terminus of the virion RNAs of EMC virus, poliovirus, rhinovirus, RAV-61 and CPMV and of mouse globin mRNA. Nucleic Acids Res. *7*, 1195–1204.

Ahlquist, P., Dasgupta, R., and Kaesberg, P. (1981). Near identity of 3- RNA secondary structure in bromoviruses and cucumber mosaic virus. Cell *23*, 183–189.

Alabi, O.J., Rwahnih, M. Al, Karthikeyan, G., Poojari, S., Fuchs, M., Rowhani, A., and Naidu, R.A. (2011). Grapevine leafroll-associated virus 1 occurs as genetically diverse populations. Phytopathology *101*, 1446–1456.

Allan, A.C., Lapidot, M., Culver, J.N., and Fluhr, R. (2001). An early tobacco mosaic virus-induced oxidative burst in tobacco indicates extracellular perception of the virus coat protein. Plant Physiol. *126*, 97–108.

Alliaume, A., Reinbold, C., Erhardt, M., Beuve, M., Hily, J.M., Lemaire, O., and Herrbach, E. (2018). Virus preparations from the mixed-infected P70 Pinot Noir accession exhibit GLRaV-1/GVA "end-to-end" particles. Arch. Virol. *163*, 3149–3154.

De Almeida, C. (2019). Interplay between uridylation and deadenylation during mRNA degradation in Arabidopsis thaliana. University of Strasbourg.

De Almeida, C., Scheer, H., Gobert, A., Fileccia, V., Martinelli, F., Zuber, H., and Gagliardi, D. (2018a). RNA uridylation and decay in plants. Philos. Trans. R. Soc. B Biol. Sci. *373*, 20180163.

De Almeida, C., Scheer, H., Zuber, H., and Gagliardi, D. (2018b). RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. Wiley Interdiscip. Rev. RNA *9*, e1440.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., et al. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science (80-.). *301*, 653–657.

Altenbach, S.B., and Howell, S.H. (1981). Identification of a satellite RNA associated with turnip crinkle virus. Virology *112*, 25–33.

Alves-Rodrigues, I., Mas, A., and Díez, J. (2007). Xenopus Xp54 and human RCK/p54 helicases functionally replace yeast Dhh1p in brome mosaic virus RNA replication. J. Virol. *81*, 4378–4380.

Amberg, D.C., Goldstein, A.L., and Cole, C.N. (1992). Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. Genes Dev. *6*, 1173–1189.

Ambros, V., and Baltimore, D. (1978). Protein is linked to the 5' end of poliovirus RNA by a phosphodiester linkage to tyrosine - PubMed. J Biol Chem *253*, 5263–5266.

Ameres, S.L., Horwich, M.D., Hung, J.H., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science *328*, 1534–1539.

Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M., et al. (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat. Struct. Mol. Biol. *20*, 1367–1376.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., and Daszak, P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol. Evol. *19*, 535–544.

Andika, I.B., Kondo, H., Nishiguchi, M., and Tamada, T. (2012). The cysteine-rich proteins of beet necrotic yellow vein virus and tobacco rattle virus contribute to efficient suppression of silencing in roots. J. Gen. Virol. *93*, 1841–1850.

Andret-Link, P., Schmitt-Keichinger, C., Demangeat, G., Komar, V., and Fuchs, M. (2004). The specific transmission of Grapevine fanleaf virus by its nematode vector Xiphinema index is solely determined by the viral coat protein. Virology *320*, 12–22.

Anindya, R., Chittori, S., and Savithri, H.S. (2005). Tyrosine 66 of Pepper vein banding virus genome-linked protein is uridylylated by RNA-dependent RNA polymerase. Virology *336*, 154–162.

Annamalai, P., and Rao, A.L.N. (2005). Dispensability of 3' tRNA-like sequence for packaging cowpea chlorotic mottle virus genomic RNAs. Virology *332*, 650–658.

Aparicio, F., Vilar, M., Perez-Payá, E., and Pallás, V. (2003). The coat protein of prunus necrotic ringspot virus specifically binds

to and regulates the conformation of its genomic RNA. Virology *313*, 213–223.

Aphasizhev, R., Suematsu, T., Zhang, L., and Aphasizheva, I. (2016). Constructive edge of uridylation-induced RNA degradation. RNA Biol. *13*, 1078–1083.

Aphasizheva, I., and Aphasizhev, R. (2016). U-Insertion/Deletion mRNA-Editing Holoenzyme: Definition in Sight. Trends Parasitol. *32*, 144–156.

Aphasizheva, I., Maslov, D., Wang, X., Huang, L., and Aphasizhev, R. (2011). Pentatricopeptide Repeat Proteins Stimulate mRNA Adenylation/Uridylation to Activate Mitochondrial Translation in Trypanosomes. Mol. Cell *42*, 106–117.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S.I., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. EMBO J. *20*, 4684–4693.

Aravind, L., and Koonin, E. V. (1999). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. *27*, 1609–1618.

Arnold, J.J., Ghosh, S.K.B., and Cameron, C.E. (1999). Poliovirus RNA-dependent RNA polymerase (3D(pol)). Divalent cation modulation of primer, template, and nucleotide selection. J. Biol. Chem. *274*, 37060–37069.

Arraiano, C.M., Mauxion, F., Viegas, S.C., Matos, R.G., and Séraphin, B. (2013). Intracellular ribonucleases involved in transcript processing and decay: precision tools for RNA. Biochim. Biophys. Acta *1829*, 491–513.

Arribere, J.A., and Fire, A.Z. (2018). Nonsense mRNA suppression via nonstop decay. Elife *7*.

Ashby, J.A., Stevenson, C.E.M., Jarvis, G.E., Lawson, D.M., and Maule, A.J. (2011). Structure-based mutational analysis of eIF4E in relation to sbm1 resistance to pea seed-borne mosaic virus in pea. PLoS One *6*.

Astuti, D., Morris, M.R., Cooper, W.N., Staals, R.H.J., Wake, N.C., Fews, G.A., Gill, H., Gentle, D., Shuib, S., Ricketts, C.J., et al. (2012). Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat. Genet. *44*, 277–284.

Aus Dem Siepen, M., Pohl, J.O., Koo, B.J., Wege, C., and Jeske, H. (2005). Poinsettia latent virus is not a cryptic virus, but a natural polerovirus–sobemovirus hybrid. Virology *336*, 240–250.

Azevedo, J., Garcia, D., Pontier, D., Ohnesorge, S., Yu, A., Garcia, S., Braun, L., Bergdoll, M., Hakimi, M.A., Lagrange, T., et al. (2010). Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Dev. *24*, 904–915.

Baccarini, A., Chauhan, H., Gardner, T.J., Jayaprakash, A.D., Sachidanandam, R., and Brown, B.D. (2011). Kinetic analysis reveals the fate of a microRNA following target regulation in mammalian cells. Curr. Biol. *21*, 369–376.

Baer, B.W., and Kornberg, R.D. (1983). The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. J. Cell Biol. *96*, 717–721.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Curr. Biol. *13*, 189–198.

Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L.A., Hart, R.P., and Kiledjian, M. (2010). Differential regulation of microRNA stability. RNA *16*, 1032–1039.

Bajczyk, M., Lange, H., Bielewicz, D., Szewc, L., Bhat, S.S., Dolata, J., Kuhn, L., Szweykowska-Kulinska, Z., Gagliardi, D., and Jarmolowski, A. (2020). SERRATE interacts with the nuclear exosome targeting (NEXT) complex to degrade primary miRNA precursors in Arabidopsis. Nucleic Acids Res. *48*, 6839–6854.

Balasubramaniam, M., Ibrahim, A., Kim, B.S., and Loesch-Fries, L.S. (2006). Arabidopsis thaliana is an asymptomatic host of Alfalfa mosaic virus. Virus Res. *121*, 215–219.

Balatsos, N.A.A., Nilsson, P., Mazza, C., Cusack, S., and Virtanen, A. (2006). Inhibition of mRNA deadenylation by the nuclear cap binding complex (CBC). J. Biol. Chem. *281*, 4517–4522.

Baltimore, D. (1971). Expression of animal virus genomes. Microbiol. Mol. Biol. Rev. *35*, 235–241.

Balzeau, J., Menezes, M.R., Cao, S., and Hagan, J.P. (2017). The LIN28/let-7 Pathway in Cancer. Front. Genet. *8*, 31.

Barajas, D., Jiang, Y., and Nagy, P.D. (2009). A unique role for the host ESCRT proteins in replication of Tomato bushy stunt virus. PLoS Pathog. *5*.

Barends, S., Bink, H.H.J., Van Den Worm, S.H.E., Pleij, C.W.A., and Kraal, B. (2003). Entrapping ribosomes for viral translation: tRNA mimicry as a molecular Trojan horse. Cell *112*, 123–129.

Barends, S., Rudinger-Thirion, J., Florentz, C., Giegé, R., Pleij, C.W.A., and Kraal, B. (2004). tRNA-like structure regulates translation of Brome mosaic virus RNA. J. Virol. *78*, 4003–4010.

Barnhart, M.D., Moon, S.L., Emch, A.W., Wilusz, C.J., and Wilusz, J. (2013). Changes in cellular mRNA stability, splicing, and polyadenylation through HuR protein sequestration by a cytoplasmic RNA virus. Cell Rep. *5*, 909–917.

Barr, J.N., Whelan, S.P.J., and Wertz, G.W. (2002). Transcriptional control of the RNA-dependent RNA polymerase of vesicular stomatitis virus. Biochim. Biophys. Acta *1577*, 337–353.

Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.M., and Heyer, W.D. (1997). A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J. Cell Biol. *136*, 761–773.

Bastin, M., and Hall, T.C. (1976). Interaction of elongation factor 1 with aminoacylated brome mosaic virus and tRNA's. J. Virol. *20*, 117–122.

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

Baumberger, N., Tsai, C.H., Lie, M., Havecker, E., and Baulcombe, D.C.C. (2007). The Polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. Curr. Biol. *17*, 1609–1614.

Beauchemin, C., and Laliberté, J.-F. (2007). The poly(A) binding protein is internalized in virus-induced vesicles or redistributed to the nucleolus during turnip mosaic virus infection. J. Virol. *81*, 10905–10913.

Beauchemin, C., Boutet, N., and Laliberté, J.-F. (2007). Visualization of the interaction between the precursors of VPg, the viral protein linked to the genome of turnip mosaic virus, and the translation eukaryotic initiation factor iso 4E in Planta. J. Virol. *81*, 775–782.

Beerens, N., Selisko, B., Ricagno, S., Imbert, I., van der Zanden, L., Snijder, E.J., and Canard, B. (2007). De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. J. Virol. *81*, 8384–8395.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev. *20*, 1885–1898.

Behrens, S.E., Tomei, L., and De Francesco, R. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. EMBO J. *15*, 12.

Beier, H., and Grimm, M. (2001). Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. Nucleic Acids Res. *29*, 4767–4782.

Beier, H., Barciszewska, M., Krupp, G., Mitnacht, R., and Gross, H.J. (1984a). UAG readthrough during TMV RNA translation: isolation and sequence of two tRNAs with suppressor activity from tobacco plants. EMBO J. *3*, 351–356.

Beier, H., Barciszewska, M., and Sickinger, H.-D. (1984b). The molecular basis for the differential translation of TMV RNA in tobacco protoplasts and wheat germ extracts. EMBO J. *3*, 1091–1096.

Beijerinck, M.W. (1968). Concerning a contagium vivum fluidum as cause of the spot disease in tobacco leaves. In Phytopathological Classics, pp. 33–52.

van Belkum, A., Cornelissen, B., Linthorst, H., Bol, J., Pley, C., and Bosch, L. (1987). tRNA-like properties of tobacco rattle virus RNA. Nucleic Acids Res. *15*, 2837–2850.

Bellott, L., Gilmer, D., and Michel, F. (2019). Hit two birds with one stone: the multiple properties of (viral) RNA silencing suppressors. Virol. (Montrouge, Fr. *23*, E38–E60.

Belostotsky, D.A. (2003). Unexpected complexity of poly(A)-binding protein gene families in flowering plants: three conserved lineages that are at least 200 million years old and possible auto- and cross-regulation. Genetics *163*, 311–319.

Belshaw, R., Gardner, A., Rambaut, A., and Pybus, O.G. (2008). Pacing a small cage: mutation and RNA viruses. Trends Ecol. Evol. *23*, 188–193.

Bernhardt, K., Linka, N., Vigelius, S.K., Weber, A.P.M., and Wiese, J. (2012). Agrobacterium-mediated Arabidopsis thaliana transformation: an overview of T-DNA binary vectors, floral dip and screening for homozygous lines. J. Endocytobiosis Cell Res. 19–28.

Bernstein, P., Peltz, S.W., and Ross, J. (1989). The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. Mol. Cell. Biol. *9*, 659–670.

Betat, H., Rammelt, C., Martin, G., and Mörl, M. (2004). Exchange of regions between bacterial poly(A) polymerase and the CCA-adding enzyme generates altered specificities. Mol. Cell *15*, 389–398.

Beuve, M., Hily, J.M., Alliaume, A., Reinbold, C., Le Maguet, J., Candresse, T., Herrbach, E., and Lemaire, O. (2018). A complex virome unveiled by deep sequencing analysis of RNAs from a French Pinot Noir grapevine exhibiting strong leafroll symptoms. Arch. Virol. *163*, 2937–2946.

Bilska, A., Kusio-Kobiałka, M., Krawczyk, P.S., Gewartowska, O., Tarkowski, B., Kobyłecki, K., Nowis, D., Golab, J., Gruchota, J., Borsuk, E., et al. (2020). Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly(A) polymerase TENT5C. Nat. Commun. *11*.

Blahna, M.T., Jones, M.R., Quinton, L.J., Matsuura, K.Y., and Mizgerd, J.P. (2011). Terminal uridyltransferase enzyme Zcchc11 promotes cell proliferation independent of its uridyltransferase activity. J. Biol. Chem. *286*, 42381–42389.

Blake, J.A., Lee, K.W., Morris, T.J., and Elthon, T.E. (2007). Effects of turnip crinkle virus infection on the structure and function of mitochondria and expression of stress proteins in turnips. Physiol. Plant. *129*, 698–706.

Blumenthal, T., and Carmichael, G.G. (1979). RNA Replication: Function and Structure of QBeta-Replicase. Annu. Rev. Biochem. *48*, 525–548.

Borden, K.L.B., and Volpon, L. (2020). The diversity, plasticity, and adaptability of cap-dependent translation initiation and the associated machinery. Https://Doi.Org/10.1080/15476286.2020.1766179 *17*, 1239–1251.

Boutrot, F., and Zipfel, C. (2017). Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. Annu. Rev. Phytopathol. *55*, 257–286.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. EMBO J. *19*, 1661–1671.

Bralley, P., and Jones, G.H. (2002). cDNA cloning confirms the polyadenylation of RNA decay intermediates in Streptomyces coelicolor. Microbiology *148*, 1421–1425.

Branscheid, A., Marchais, A., Schott, G., Lange, H., Gagliardi, D., Andersen, S.U., Voinnet, O., and Brodersen, P. (2015). SKI2

mediates degradation of RISC 5'-cleavage fragments and prevents secondary siRNA production from miRNA targets in Arabidopsis. Nucleic Acids Res. *43*, 10975–10988.

Briani, F., Carzaniga, T., and Dehò, G. (2016). Regulation and functions of bacterial PNPase. Wiley Interdiscip. Rev. RNA *7*, 241–258.

Briggs, M.W., Burkard, K.T.D., and Butler, J.S. (1998). Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J. Biol. Chem. *273*, 13255–13263.

Broadbent, L., and Heathcote, G.D. (1958). Properties and host range of turnip crinkle, rosette and yellow mosaic viruses. Ann. Appl. Biol. *46*, 585–592.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. Science *320*, 1185–1190.

Brown, J.T., Bai, X., and Johnson, A.W. (2000). The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo. RNA *6*, 449–457.

Bujarski, J.J., and Kaesberg, P. (1986). Genetic recombination between RNA components of a multipartite plant virus. Nature *321*, 528–531.

Bujarski, J., Gallitelli, D., García-Arenal, F., Pallás, V., Palukaitis, P., Krishna Reddy, M., and Wang, A. (2019). ICTV Virus Taxonomy Profile: Bromoviridae. J. Gen. Virol. *100*, 1206–1207.

Burkard, K.T.D., and Butler, J.S. (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Mol. Cell. Biol. *20*, 604–616.

Buschauer, R., Matsuo, Y., Sugiyama, T., Chen, Y.H., Alhusaini, N., Sweet, T., Ikeuchi, K., Cheng, J., Matsuki, Y., Nobuta, R., et al. (2020). The Ccr4-Not complex monitors the translating ribosome for codon optimality. Science *368*.

Callahan, K.P., and Butler, J.S. (2008). Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. Nucleic Acids Res. *36*, 6645–6655.

Campos-Guillén, J., Bralley, P., Jones, G.H., Bechhofer, D.H., and Olmedo-Alvarez, G. (2005). Addition of poly(A) and heteropolymeric 3' ends in Bacillus subtilis wild-type and polynucleotide phosphorylase-deficient strains. J. Bacteriol. *187*, 4698–4706.

Caponigro, G., and Parker, R. (1995). Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. Genes Dev. *9*, 2421–2432.

Careno, D.A., Santangelo, S.P., Macknight, R.C., and Yanovsky, M.J. (2022). The 5′-3′ exoribonuclease XRN4 modulates the plant circadian network in Arabidopsis. BioRxiv 2022.07.06.499002.

Carpenter, C.D., and Simon, A.E. (1996). In vivo restoration of biologically active 3' ends of virus-associated RNAs by nonhomologous RNA recombination and replacement of a terminal motif. J. Virol. *70*, 478–486.

Carpenter, C.D., Oh, J.W., Zhang, C., and Simon, A.E. (1995). Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. J. Mol. Biol. *245*, 608–622.

Carpentier, M.C., Deragon, J.M., Jean, V., Seng Hour Vichet, B., Bousquet-Antonelli, C., and Merret, R. (2020). Monitoring of XRN4 Targets Reveals the Importance of Cotranslational Decay during Arabidopsis Development. Plant Physiol. *184*, 1251–1262.

Carrington, J.C., Morris, T.J., Stockley, P.G., and Harrison, S.C. (1987). Structure and assembly of turnip crinkle virus. IV. Analysis of the coat protein gene and implications of the subunit primary structure. J. Mol. Biol. *194*, 265–276.

Carrington, J.C., Heaton, L.A., Zuidema, D., Hillman, B.I., and Morris, T.J. (1989). The genome structure of turnip crinkle virus. Virology *170*, 219–226.

Carroll, T.W. (1970). Relation of barley stripe mosaic virus to plastids. Virology *42*, 1015–1022.

Castaño, A., Ruiz, L., and Hernández, C. (2009). Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus. Virology *386*, 417–426.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014). TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3′ End Modifications. Mol. Cell *53*, 1044–1052.

Chang, H.M., Triboulet, R., Thornton, J.E., and Gregory, R.I. (2013). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature *497*, 244–248.

Chantarachot, T., Sorenson, R.S., Hummel, M., Ke, H., Kettenburg, A.T., Chen, D., Aiyetiwa, K., Dehesh, K., Eulgem, T., Sieburth, L.E., et al. (2020). DHH1/DDX6-like RNA helicases maintain ephemeral half-lives of stress-response mRNAs. Nat. Plants *6*, 675–685.

Chapman, R.F. (1998). The Insects: Structure and Function. Insects.

Chapman, M.R., and Kao, C.C. (1999). A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. J. Mol. Biol. *286*, 709–720.

Chapman, E.G., Moon, S.L., Wilusz, J., and Kieft, J.S. (2014). RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA. Elife *3*.

Chatterjee, S., and Großhans, H. (2009). Active turnover modulates mature microRNA activity in Caenorhabditis elegans. Nature *461*, 546–549.

Chen, M.-H., and Frey, T.K. (1999). Mutagenic analysis of the 3' cis-acting elements of the rubella virus genome. J. Virol. *73*,

3386–3403.

Chen, I.-H., Chou, W.-J., Lee, P.-Y., Hsu, Y.-H., and Tsai, C.-H. (2005). The AAUAAA Motif of Bamboo Mosaic Virus RNA Is Involved in Minus-Strand RNA Synthesis and Plus-Strand RNA Polyadenylation . J. Virol. *79*, 14555–14561.

Chen, I.H., Cheng, J.H., Huang, Y.W., Lin, N.S., Hsu, Y.H., and Tsai, C.H. (2013). Characterization of the polyadenylation activity in a replicase complex from Bamboo mosaic virus-infected Nicotiana benthamiana plants. Virology *444*, 64–70.

Chen, Z., Li, Y., and Krug, R.M. (1999). Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. EMBO J. *18*, 2273–2283.

Cheng, X., and Wang, A. (2016). The Potyvirus Silencing Suppressor Protein VPg Mediates Degradation of SGS3 via Ubiquitination and Autophagy Pathways. J. Virol. *91*.

Cheng, C.P., Jaag, H.M., Jonczyk, M., Serviene, E., and Nagy, P.D. (2007). Expression of the Arabidopsis Xrn4p 5′–3′ exoribonuclease facilitates degradation of tombusvirus RNA and promotes rapid emergence of viral variants in plants. Virology *368*, 238–248.

Cheng, J.-H., Peng, C.-W., Hsu, Y.-H., and Tsai, C.-H. (2002). The synthesis of minus-strand RNA of bamboo mosaic potexvirus initiates from multiple sites within the poly(A) tail. J. Virol. *76*, 6114–6120.

Cheng, X., Xiong, R., Li, Y., Li, F., Zhou, X., and Wang, A. (2017). Sumoylation of Turnip mosaic virus RNA Polymerase Promotes Viral Infection by Counteracting the Host NPR1-Mediated Immune Response. Plant Cell *29*, 508–525.

Chenon, M., Camborde, L., Cheminant, S., and Jupin, I. (2012). A viral deubiquitylating enzyme targets viral RNA-dependent RNA polymerase and affects viral infectivity. EMBO J. *31*, 741–753.

Cheo, P.C. (1970). Subliminal Infection of Cotton by Tobacco Mosaic Virus. Phytopathology *60*, 41.

Cheo, P.C. (1971). Effect in different plant species of continuous light and dark treatment on tobacco mosaic virus replicating capacity. Virology *46*, 256–265.

Chiba, S., Hleibieh, K., Delbianco, A., Klein, E., Ratti, C., Ziegler-Graff, V., Bouzoubaa, S., and Gilmer, D. (2013). The Benyvirus RNA Silencing Suppressor Is Essential for Long-Distance Movement, Requires Both Zinc-Finger and NoLS Basic Residues but Not a Nucleolar Localization for Its Silencing-Suppression Activity. MPMI *26*, 168–181.

Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., Van Erp, H., and Green, P.J. (2004). AtPARN is an essential poly(A) ribonuclease in Arabidopsis. Gene *328*, 95–102.

Chisholm, J., Zhang, G., Wang, A., and Sanfaçon, H. (2007). Peripheral association of a polyprotein precursor form of the RNA-dependent RNA polymerase of Tomato ringspot virus with the membrane-bound viral replication complex. Virology *368*, 133–144.

Chisholm, S.T., Coaker, G., Day, B., and Staskawicz, B.J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. Cell *124*, 803–814.

Chiu, M.H., Chen, I.H., Baulcombe, D.C., and Tsai, C.H. (2010). The silencing suppressor P25 of Potato virus X interacts with Argonaute1 and mediates its degradation through the proteasome pathway. Mol. Plant Pathol. *11*, 641–649.

Chiu, S.Y., Lejeune, F., Ranganathan, A.C., and Maquat, L.E. (2004). The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex. Genes Dev. *18*, 745–754.

Chlebowski, A., Tomecki, R., López, M.E.G., Séraphin, B., and Dziembowski, A. (2011). Catalytic properties of the eukaryotic exosome. Adv. Exp. Med. Biol. *702*, 63–78.

Cho, T.J., and Dreher, T.W. (2006). Encapsidation of genomic but not subgenomic Turnip yellow mosaic virus RNA by coat protein provided in trans. Virology *356*, 126–135.

Choe, J., Oh, N., Park, S., Lee, Y.K., Song, O.K., Locker, N., Chi, S.G., and Kim, Y.K. (2012). Translation initiation on mRNAs bound by nuclear cap-binding protein complex CBP80/20 requires interaction between CBP80/20-dependent translation initiation factor and eukaryotic translation initiation factor 3g. J. Biol. Chem. *287*, 18500–18509.

Choe, J., Ryu, I., Park, O.H., Park, J., Cho, H., Yoo, J.S., Chi, S.W., Kim, M.K., Song, H.K., and Kim, Y.K. (2014). eIF4AIII enhances translation of nuclear cap-binding complex-bound mRNAs by promoting disruption of secondary structures in 5'UTR. Proc. Natl. Acad. Sci. U. S. A. *111*, E4577–E4586.

Choi, Y.G., Dreher, T.W., and Rao, A.L.N. (2002). tRNA elements mediate the assembly of an icosahedral RNA virus. Proc. Natl. Acad. Sci. U. S. A. *99*, 655–660.

Chowdhury, S.R., and Savithri, H.S. (2011). Interaction of Sesbania mosaic virus movement protein with VPg and P10: implication to specificity of genome recognition. PLoS One *6*.

Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. RNA *13*, 998–1016.

Chung, B.Y.W., Miller, W.A., Atkins, J.F., and Firth, A.E. (2008). An overlapping essential gene in the Potyviridae. Proc. Natl. Acad. Sci. U. S. A. *105*, 5897–5902.

Cimino, P.A., Nicholson, B.L., Wu, B., Xu, W., and White, K.A. (2011). Multifaceted regulation of translational readthrough by RNA replication elements in a tombusvirus. PLoS Pathog. *7*.

Citores, L., Iglesias, R., and Ferreras, J.M. (2021). Antiviral Activity of Ribosome-Inactivating Proteins. Toxins (Basel). *13*.

Cohen, L.S., Mikhli, C., Jiao, X., Kiledjian, M., Kunkel, G., and Davis, R.E. (2005). Dcp2 Decaps m2,2,7GpppN-capped RNAs, and its activity is sequence and context dependent. Mol. Cell. Biol. *25*, 8779–8791.

Cohen, Y., Qu, F., Gisel, A., Morris, T.J., and Zambryski, P.C. (2000). Nuclear localization of turnip crinkle virus movement protein p8. Virology *273*, 276–285.

Coll, N.S., Epple, P., and Dangl, J.L. (2011). Programmed cell death in the plant immune system. Cell Death Differ. *18*, 1247–1256.

Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip. Rev. RNA *7*, 438–454.

Conrad, N.K., Mili, S., Marshall, E.L., Shu, M. Di, and Steitz, J.A. (2006). Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. Mol. Cell *24*, 943–953.

Copeland, P.R., and Wormington, M. (2001). The mechanism and regulation of deadenylation: Identification and characterization of Xenopus PARN. RNA *7*, 875–886.

da Costa, P.J., Menezes, J., Saramago, M., García-Moreno, J.F., Santos, H.A., Gama-Carvalho, M., Arraiano, C.M., Viegas, S.C., and Romão, L. (2019). Experimental supporting data on DIS3L2 over nonsense-mediated mRNA decay targets in human cells. Data Br. *28*.

Cotton, S., Grangeon, R., Thivierge, K., Mathieu, I., Ide, C., Wei, T., Wang, A., and Laliberté, J.-F. (2009). Turnip mosaic virus RNA replication complex vesicles are mobile, align with microfilaments, and are each derived from a single viral genome. J. Virol. *83*, 10460–10471.

Couto, D., and Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. Nat. Rev. Immunol. *16*, 537–552.

Cranston, P.S., and Gullan, P.J. (2009). Phylogeny of Insects. Encycl. Insects 780–793.

Cross, S.T., Michalski, D., Miller, M.R., and Wilusz, J. (2019). RNA regulatory processes in RNA virus biology. Wiley Interdiscip. Rev. RNA *10*.

Cui, H., Tsuda, K., and Parker, J.E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. Annu. Rev. Plant Biol. *66*, 487–511.

Czesnick, H., and Lenhard, M. (2016). Antagonistic control of flowering time by functionally specialized poly(A) polymerases in Arabidopsis thaliana. Plant J. *88*, 570–583.

D'Alonzo, M., Delbianco, A., Lanzoni, C., Autonell, C.R., Gilmer, D., and Ratti, C. (2012). Beet soil-borne mosaic virus RNA-4 encodes a 32 kDa protein involved in symptom expression and in virus transmission through Polymyxa betae. Virology *423*, 187–194.

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature *468*, 452–456.

Danthinne, X., Seurinck, J., Meulewaeter, F., Van Montagu, M., and Cornelissen, M. (1993). The 3' untranslated region of satellite tobacco necrosis virus RNA stimulates translation in vitro. Mol. Cell. Biol. *13*, 3340–3349.

Dawson, W.O., and Hilf, M.E. (1992). HOST-RANGE DETERMINANTS OF PLANT VIRUSES. Annu. Rev. Plant Physiol. Plant Mol. Bioi *43*, 527–555.

Decroly, E., Ferron, F., Lescar, J., and Canard, B. (2011). Conventional and unconventional mechanisms for capping viral mRNA. Nat. Rev. Microbiol. *10*, 51–65.

DeFalco, T.A., and Zipfel, C. (2021). Molecular mechanisms of early plant pattern-triggered immune signaling. Mol. Cell *81*, 3449–3467.

Dehlin, E., Wormington, M., Körner, C.G., and Wahle, E. (2000). Cap-dependent deadenylation of mRNA. EMBO J. *19*, 1079.

Deiman, B.A.L.M., Koenen, A.K., Verlaan, P.W.G., and Pleij, C.W.A. (1998). Minimal template requirements for initiation of minus-strand synthesis in vitro by the RNA-dependent RNA polymerase of turnip yellow mosaic virus. J. Virol. *72*, 3965–3972.

Delan-Forino, C., Spanos, C., Rappsilber, J., and Tollervey, D. (2020). Substrate specificity of the TRAMP nuclear surveillance complexes. Nat. Commun. *11*.

Delis, C., Krokida, A., Tomatsidou, A., Tsikou, D., Beta, R.A.A., Tsioumpekou, M., Moustaka, J., Stravodimos, G., Leonidas, D.D., Balatsos, N.A.A., et al. (2016). AtHESPERIN: a novel regulator of circadian rhythms with poly(A)-degrading activity in plants. RNA Biol. *13*, 68–82.

Depledge, D.P., and Wilson, A.C. (2020). Using Direct RNA Nanopore Sequencing to Deconvolute Viral Transcriptomes. Curr. Protoc. Microbiol. *57*.

Derrien, B., Baumberger, N., Schepetilnikov, M., Viotti, C., De Cillia, J., Ziegler-Graff, V., Isono, E., Schumacher, K., and Genschik, P. (2012). Degradation of the antiviral component ARGONAUTE1 by the autophagy pathway. Proc. Natl. Acad. Sci. U. S. A. *109*, 15942–15946.

Deyholos, M.K., Cavaness, G.F., Hall, B., King, E., Punwani, J., Van Norman, J., and Sieburth, L.E. (2003). VARICOSE, a WD-domain protein, is required for leaf blade development. Development *130*, 6577–6588.

Dickson, A.M., Anderson, J.R., Barnhart, M.D., Sokoloski, K.J., Oko, L., Opyrchal, M., Galanis, E., Wilusz, C.J., Morrison, T.E., and Wilusz, J. (2012). Dephosphorylation of HuR protein during alphavirus infection is associated with HuR relocalization to the cytoplasm. J. Biol. Chem. *287*, 36229–36238.

Díez, J., Ishikawa, M., Kaido, M., and Ahlquist, P. (2000). Identification and characterization of a host protein required for efficient template selection in viral RNA replication. Proc. Natl. Acad. Sci. U. S. A. *97*, 3913–3918.

Van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., and Séraphin, B. (2002). Human Dcp2: a catalytically active mRNA

decapping enzyme located in specific cytoplasmic structures. EMBO J. *21*, 6915–6924.

Dilweg, I.W., Gultyaev, A.P., and Olsthoorn, R.C. (2019). Structural features of an Xrn1-resistant plant virus RNA. RNA Biol. *16*, 838.

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

Ding, X.S., Liu, J., Cheng, N.H., Folimonov, A., Hou, Y.M., Bao, Y., Katagi, C., Carter, S.A., and Nelson, R.S. (2004). The Tobacco mosaic virus 126-kDa protein associated with virus replication and movement suppresses RNA silencing. Mol. Plant. Microbe. Interact. *17*, 583–592.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561–564.

Domashevskiy, A. V., Williams, S., Kluge, C., and Cheng, S.Y. (2017). Plant Translation Initiation Complex eIFiso4F Directs Pokeweed Antiviral Protein to Selectively Depurinate Uncapped Tobacco Etch Virus RNA. Biochemistry *56*, 5980–5990.

Domingues, M.N., Sforça, M.L., Soprano, A.S., Lee, J., De Souza, T.D.A.C.B., Cassago, A., Portugal, R.V., De Mattos Zeri, A.C., Murakami, M.T., Sadanandom, A., et al. (2015). Structure and Mechanism of Dimer–Monomer Transition of a Plant Poly(A)-Binding Protein upon RNA Interaction: Insights into Its Poly(A) Tail Assembly. J. Mol. Biol. *427*, 2491–2506.

Dorcey, E., Rodriguez-Villalon, A., Salinas, P., Santuari, L., Pradervand, S., Harshman, K., and Hardtke, C.S. (2012). Context-dependent dual role of SKI8 homologs in mRNA synthesis and turnover. PLoS Genet. *8*.

Dreher, T.W. (2010). Viral tRNAs and tRNA-like structures. Wiley Interdiscip. Rev. RNA *1*, 402–414.

Dreher, T.W., and Goodwin, J.B. (1998). Transfer RNA mimicry among tymoviral genomic RNAs ranges from highly efficient to vestigial. Nucleic Acids Res. *26*, 4356–4364.

Dreher, T.W., and Hall, T.C. (1988). Mutational analysis of the tRNA mimicry of brome mosaic virus RNA. Sequence and structural requirements for aminoacylation and 3'-adenylation. J. Mol. Biol. *201*, 41–55.

Dreher, T.W., Bujarski, J.J., and Hall, T.C. (1984). Mutant viral RNAs synthesized in vitro show altered aminoacylation and replicase template activities. Nature *311*, 171–175.

Dreher, T.W., Edwards, M.C., Gibbs, A.J., Haenni, A.-L., Hammond, R.W., Jupin, I., Koenig, R., Sabanadzovic, S., and Martelli, G.P. (2009). ICTV Virus Taxonomy Profile: Tymoviridae 2009.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun. *7*.

Dufresne, P.J., Ubalijoro, E., Fortin, M.G., and Laliberte, J.F. (2008a). Arabidopsis thaliana class II poly(A)-binding proteins are required for efficient multiplication of turnip mosaic virus. J. Gen. Virol. *89*, 2339–2348.

Dufresne, P.J., Thivierge, K., Cotton, S., Beauchemin, C., Ide, C., Ubalijoro, E., Laliberté, J.F., and Fortin, M.G. (2008b). Heat shock 70 protein interaction with Turnip mosaic virus RNA-dependent RNA polymerase within virus-induced membrane vesicles. Virology *374*, 217–227.

Dumas, P., Moras, D., Florentz, C., Giege, R., Verlaan, P., Van Belkum, A., and Pleij, C.W.A. (1987). 3-D graphics modelling of the tRNA-like 3'-end of turnip yellow mosaic virus RNA: structural and functional implications. J. Biomol. Struct. Dyn. *4*, 707–728.

Dunckley, T., and Parker, R. (1999). The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. EMBO J. *18*, 5411–5422.

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K.S., and Robaglia, C. (2002). The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. Plant J. *32*, 927–934.

Dupressoir, A., Morel, A.P., Barbot, W., Loireau, M.P., Corbo, L., and Heidmann, T. (2001). Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. BMC Genomics *2*.

Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat. Struct. Mol. Biol. *14*, 15–22.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Edwardson, J.R., and Christie, R.G. (1986). Viruses Infecting Forage Legumes.

Eggen, R., Verver, J., Wellink, J., De Jong, A., Goldbach, R., and van Kammen, A. (1989). Improvements of the infectivity of in vitro transcripts from cloned cowpea mosaic virus cDNA: impact of terminal nucleotide sequences. Virology *173*, 447–455.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., et al. (2009). Real-time DNA sequencing from single polymerase molecules. Science (80-.). *323*, 133–138.

Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., and Bartel, D.P. (2020). The Dynamics of Cytoplasmic mRNA Metabolism. Mol. Cell *77*, 786-799.e10.

Estrella, M.A., Du, J., and Korennykh, A. (2018). Crystal Structure of Human Nocturnin Catalytic Domain. Sci. Reports 2018 81 *8*, 1–8.

Estrella, M.A., Du, J., Chen, L., Rath, S., Prangley, E., Chitrakar, A., Aoki, T., Schedl, P., Rabinowitz, J., and Korennykh, A. (2019). The metabolites NADP+ and NADPH are the targets of the circadian protein Nocturnin (Curled). Nat. Commun. 2019 101 *10*, 1–10.

Fabian, M.R., Na, H., Ray, D., and White, K.A. (2003). 3′-Terminal RNA secondary structures are important for accumulation of tomato bushy stunt virus DI RNAs. Virology *313*, 567–580.

Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., et al. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nat. Struct. Mol. Biol. *18*, 1211–1217.

Faehnle, C.R., Walleshauser, J., and Joshua-Tor, L. (2014). Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. Nature *514*, 252–256.

Faehnle, C.R., Walleshauser, J., and Joshua-Tor, L. (2017). Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. Nat. Struct. Mol. Biol. *24*, 658–665.

Fan, X., Hong, N., Dong, Y., Ma, Y., Zhang, Z.P., Ren, F., Hu, G., Zhou, J., and Wang, G. (2015). Genetic diversity and recombination analysis of grapevine leafroll-associated virus 1 from China. Arch. Virol. *160*, 1669–1678.

Fechter, P., Giegé, R., and Rudinger-Thirion, J. (2001). Specific tyrosylation of the bulky tRNA-like structure of brome mosaic virus RNA relies solely on identity nucleotides present in its amino acid-accepting domain. J. Mol. Biol. *309*, 387–399.

Felden, B., Florentz, C., Mcpherson, A., and Giegé, R. (1994). A histidine accepting tRNA-like fold at the 3'-end of satellite tobacco mosaic virus RNA. Nucleic Acids Res. *22*, 2882–2886.

de Felippes, F.F., and Waterhouse, P.M. (2020). The Whys and Wherefores of Transitivity in Plants. Front. Plant Sci. *11*.

Fellers, J., Wan, J., Hong, Y., Collins, G.B., and Hunt, A.G. (1998). In vitro interactions between a potyvirus-encoded, genome-linked protein and RNA-dependent RNA polymerase. J. Gen. Virol. *79 (Pt 8)*, 2043–2049.

Ferrier, E. (2013). Rôle et mode d'action de l'UTP : RNA Uridylyltransférase URT1 dans l'uridylation et la dégradation des ARNm chez Aradopsis thaliana. Université de Strasbourg.

Flobinus, A., Chevigny, N., Charley, P.A., Seissler, T., Klein, E., Bleykasten-Grosshans, C., Ratti, C., Bouzoubaa, S., Wilusz, J., and Gilmer, D. (2018). Beet Necrotic Yellow Vein Virus Noncoding RNA Production Depends on a 5'→3' Xrn Exoribonuclease Activity. Viruses *10*.

Florentz, C., Briand, J.P., Romby, P., Hirth, L., Ebel, J.P., and Glegé, R. (1982). The tRNA-like structure of turnip yellow mosaic virus RNA: structural organization of the last 159 nucleotides from the 3′ OH terminus. EMBO J. *1*, 269–276.

Fortes, P., Inada, T., Preiss, T., Hentze, M.W., Mattaj, I.W., and Sachs, A.B. (2000). The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Mol. Cell *6*, 191–196.

Fresco, L.D., and Buratowski, S. (1996). Conditional mutants of the yeast mRNA capping enzyme show that the cap enhances, but is not required for, mRNA splicing. RNA *2*, 584–596.

Fromm, S.A., Truffault, V., Kamenz, J., Braun, J.E., Hoffmann, N.A., Izaurralde, E., and Sprangers, R. (2012). The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex. EMBO J. *31*, 279–290.

Fromont-Racine, M., Mayes, A.E., Brunet-Simon, A., Rain, J.C., Colley, A., Dix, I., Decourty, L., Joly, N., Ricard, F., Beggs, J.D., et al. (2000). Genome-Wide Protein Interaction Screens Reveal Functional Networks Involving Sm-Like Proteins. Yeast *17*, 95.

Fu, Z.Q., and Dong, X. (2013). Systemic acquired resistance: turning local infection into global defense. Annu. Rev. Plant Biol. *64*, 839–863.

Fuchs, M., Schmitt-Keichinger, C., and Sanfaçon, H. (2017). A Renaissance in Nepovirus Research Provides New Insights Into Their Molecular Interface With Hosts and Vectors. Adv. Virus Res. *97*, 61–105.

Fuchs, M., Bar-Joseph, M., Candresse, T., Maree, H.J., Martelli, G.P., Melzer, M.J., Menzel, W., Minafra, A., and Sabanadzovic, S. (2020). ICTV Virus Taxonomy Profile: Closteroviridae. J. Gen. Virol. *101*, 364–365.

Fullerton, S.W.B., Blaschke, M., Coutard, B., Gebhardt, J., Gorbalenya, A., Canard, B., Tucker, P.A., and Rohayem, J. (2007). Structural and functional characterization of sapovirus RNA-dependent RNA polymerase. J. Virol. *81*, 1858–1871.

Funakoshi, Y., Doi, Y., Hosoda, N., Uchida, N., Osawa, M., Shimada, I., Tsujimoto, M., Suzuki, T., Katada, T., and Hoshino, S.I. (2007). Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. Genes Dev. *21*, 3135–3148.

Funk, A., Truong, K., Nagasaki, T., Torres, S., Floden, N., Balmori Melian, E., Edmonds, J., Dong, H., Shi, P.-Y., and Khromykh, A.A. (2010). RNA structures required for production of subgenomic flavivirus RNA. J. Virol. *84*, 11407–11417.

Furuichi, Y., and Miura, K.I. (1975). A blocked structure at the 5' terminus of mRNA from cytoplasmic polyhedrosis virus. Nature *253*, 374–375.

Fütterer, J., and Hohn, T. (1996). Translation in plants-rules and exceptions. Plant Mol. Biol. 1996 321 *32*, 159–189.

Gadh, I.P.S., and Hari, V. (1986). Association of tobacco etch virus related RNA with chloroplasts in extracts of infected plants. Virology *150*, 304–307.

Gaire, F., Schmitt, C., Stussi-Garaud, C., Pinck, L., and Ritzenthaler, C. (1999). Protein 2A of Grapevine Fanleaf Nepovirus Is Implicated in RNA2 Replication and Colocalizes to the Replication Site. Virology *264*, 25–36.

Galão, R.P., Chari, A., Alves-Rodrigues, I., Lobão, D., Mas, A., Kambach, C., Fischer, U., and Díez, J. (2010). LSm1-7 complexes bind to specific sites in viral RNA genomes and regulate their translation and replication. RNA *16*, 817–827.

Gallois, J.L., Charron, C., Sanchez, F., Pagny, G., Houvenaghel, M.C., Moretti, A., Ponz, F., Revers, F., Caranta, C., and German-Retana, S. (2010). Single amino acid changes in the turnip mosaic virus viral genome-linked protein (VPg) confer virulence towards Arabidopsis thaliana mutants knocked out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G. J. Gen. Virol. *91*,

288–293.

Garcia, D., Garcia, S., and Voinnet, O. (2014). Nonsense-mediated decay serves as a general viral restriction mechanism in plants. Cell Host Microbe *16*, 391–402.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575–578.

Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A Link Between mRNA Turnover and RNA Interference in Arabidopsis. Science (80-.). *306*, 1046–1048.

Geng, G., Yu, C., Li, X., and Yuan, X. (2019). Variable 3'polyadenylation of Wheat yellow mosaic virus and its novel effects on translation and replication. Virol. J. *16*, 23.

Geng, G., Wang, D., Liu, Z., Wang, Y., Zhu, M., Cao, X., Yu, C., and Yuan, X. (2021). Translation of Plant RNA Viruses. Viruses *13*.

Gewartowska, O., Aranaz-Novaliches, G., Krawczyk, P.S., Mroczek, S., Kusio-Kobiałka, M., Tarkowski, B., Spoutil, F., Benada, O., Kofroňová, O., Szwedziak, P., et al. (2021). Cytoplasmic polyadenylation by TENT5A is required for proper bone formation. Cell Rep. *35*.

Ghosh, A., and Lima, C.D. (2010). Enzymology of RNA cap synthesis. Wiley Interdiscip. Rev. RNA *1*, 152–172.

Ghoshal, B., and Sanfaçon, H. (2014). Temperature-dependent symptom recovery in Nicotiana benthamiana plants infected with tomato ringspot virus is associated with reduced translation of viral RNA2 and requires ARGONAUTE 1. Virology *456–457*, 188–197.

Giampetruzzi, A., Roumi, V., Roberto, R., Malossini, U., Yoshikawa, N., La Notte, P., Terlizzi, F., Credi, R., and Saldarelli, P. (2012). A new grapevine virus discovered by deep sequencing of virus- and viroid-derived small RNAs in Cv Pinot gris. Virus Res. *163*, 262–268.

Giegé, R., Briand, J. -P, Mengual, R., Ebel, J. -P, and Hirth, L. (1978). Valylation of the two RNA components of turnip-yellow mosaic virus and specificity of the tRNA aminoacylation reaction. Eur. J. Biochem. *84*, 251–256.

Gilmer, D., and Ratti, C. (2017). ICTV Virus Taxonomy Profile: Benyviridae. J. Gen. Virol. *98*, 1571.

Girard, C., Verheggen, C., Neel, H., Cammas, A., Vagner, S., Soret, J., Bertrand, E., and Bordonné, R. (2008). Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing. J. Biol. Chem. *283*, 2060–2069.

Godwin, A.R., Kojima, S., Green, C.B., and Wilusz, J. (2013). Kiss your tail goodbye: the role of PARN, Nocturnin, and Angel deadenylases in mRNA biology. Biochim. Biophys. Acta *1829*, 571.

Golisz, A., Sikorski, P.J., Kruszka, K., and Kufel, J. (2013). Arabidopsis thaliana LSM proteins function in mRNA splicing and degradation. Nucleic Acids Res. *41*, 6232–6249.

Goodwin, J.B., and Dreher, T.W. (1998). Transfer RNA mimicry in a new group of positive-strand RNA plant viruses, the furoviruses: differential aminoacylation between the RNA components of one genome. Virology *246*, 170–178.

Görlach, M., Burd, C.G., and Dreyfuss, G. (1994). The mRNA poly(A)-binding protein: localization, abundance, and RNA-binding specificity. Exp. Cell Res. *211*, 400–407.

Goss, D.J., and Kleiman, F.E. (2013). Poly(A) binding proteins: are they all created equal? Wiley Interdiscip. Rev. RNA *4*, 167–179.

Grangeon, R., Cotton, S., and Laliberté, J.-F. (2010). A model for the biogenesis of Turnip mosaic virus replication factories. Commun. Integr. Biol. *3*, 363–365.

Grangeon, R., Jiang, J., Wan, J., Agbeci, M., Zheng, H., and Laliberté, J.F. (2013). 6K2-induced vesicles can move cell to cell during turnip mosaic virus infection. Front. Microbiol. *4*.

Green, C.B., and Besharse, J.C. (1996). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. Proc. Natl. Acad. Sci. U. S. A. *93*, 14884–14888.

Grzela, R., Szolajska, E., Ebel, C., Madern, D., Favier, A., Wojtal, I., Zagorski, W., and Chroboczek, J. (2008). Virulence factor of potato virus Y, genome-attached terminal protein VPg, is a highly disordered protein. J. Biol. Chem. *283*, 213–221.

Guan, H., and Simon, A.E. (2000). Polymerization of nontemplate bases before transcription initiation at the 3' ends of templates by an RNA-dependent RNA polymerase: An activity involved in 3' end repair of viral RNAs. Proc. Natl. Acad. Sci. U. S. A. *97*, 12451–12456.

Gudipati, R.K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L.M., Jacquier, A., and Libri, D. (2012). Extensive Degradation of RNA Precursors by the Exosome in Wild-Type Cells. Mol. Cell *48*, 409–421.

Guilford, P.J., Beck, D.L., and Forster, R.L.S. (1991). Influence of the poly(A) tail and putative polyadenylation signal on the infectivity of white clover mosaic potexvirus. Virology *182*, 61–67.

Gunawardene, C.D., Newburn, L.R., and Andrew White, K. (2019). A 212-nt long RNA structure in the Tobacco necrosis virus-D RNA genome is resistant to Xrn degradation. Nucleic Acids Res. *47*, 9329–9342.

Gunawardene, C.D., Im, J.S.H., and White, K.A. (2021). RNA Structure Protects the 5' End of an Uncapped Tombusvirus RNA Genome from Xrn Digestion. J. Virol. *95*.

Gutiérrez-Vázquez, C., Enright, A.J., Rodríguez-Galán, A., Pérez-García, A., Collier, P., Jones, M.R., Benes, V., Mizgerd, J.P., Mittelbrunn, M., Ramiro, A.R., et al. (2017). 3' Uridylation controls mature microRNA turnover during CD4 T-cell activation. RNA

23, 882–891.

Guydosh, N.R., Kimmig, P., Walter, P., and Green, R. (2017). Regulated Ire1-dependent mRNA decay requires no-go mRNA degradation to maintain endoplasmic reticulum homeostasis in S. pombe. Elife *6*.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 Are Endogenous RNA Silencing Suppressors. Plant Cell *19*, 3451–3461.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat. Rev. Mol. Cell Biol. *15*, 509–524.

Haas, G., Braun, J.E., Igreja, C., Tritschler, F., Nishihara, T., and Izaurralde, E. (2010). HPat provides a link between deadenylation and decapping in metazoa. J. Cell Biol. *189*, 289–302.

Haas, G., Cetin, S., Messmer, M., Chane-Woon-Ming, B., Terenzi, O., Chicher, J., Kuhn, L., Hammann, P., and Pfeffer, S. (2016). Identification of factors involved in target RNA-directed microRNA degradation. Nucleic Acids Res. *44*, 2873–2887.

Hacker, D.L., and Sivakumaran, K. (1997). Mapping and Expression of Southern Bean Mosaic Virus Genomic and Subgenomic RNAs. Virology *234*, 317–327.

Hacker, D.L., Petty, I.T.D., Wei, N., and Morris, T.J. (1992). Turnip crinkle virus genes required for RNA replication and virus movement. Virology *186*, 1–8.

Hafrén, A., Lõhmus, A., and Mäkinen, K. (2015). Formation of Potato Virus A-Induced RNA Granules and Viral Translation Are Interrelated Processes Required for Optimal Virus Accumulation. PLoS Pathog. *11*.

Hafrén, A., Üstün, S., Hochmuth, A., Svenning, S., Johansen, T., and Hofius, D. (2018). Turnip Mosaic Virus Counteracts Selective Autophagy of the Viral Silencing Suppressor HCpro. Plant Physiol. *176*, 649–662.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat. Struct. Mol. Biol. *16*, 1021–1025.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000–1011.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The Yeast Ski Complex: Crystal Structure and RNA Channeling to the Exosome Complex. Cell *154*, 814–826.

Halstead, J.M., Lionnet, T., Wilbertz, J.H., Wippich, F., Ephrussi, A., Singer, R.H., and Chao, J.A. (2015). Translation. An RNA biosensor for imaging the first round of translation from single cells to living animals. Science *347*, 1367–1370.

Harnisch, C., Cuzic-Feltens, S., Dohm, J.C., Götze, M., Himmelbauer, H., and Wahle, E. (2016). Oligoadenylation of 3' decay intermediates promotes cytoplasmic mRNA degradation in Drosophila cells. RNA *22*, 428–442.

Harrison, B., and Zimmerman, S.B. (1984). Polymer-stimulated ligation: enhanced ligation of oligo- and polynucleotides by T4 RNA ligase in polymer solutions. Nucleic Acids Res. *12*, 8235.

Hausmann, S., Garcin, D., Delenda, C., and Kolakofsky, D. (1999). The versatility of paramyxovirus RNA polymerase stuttering. J. Virol. *73*, 5568–5576.

Haynes, C., Oldfield, C.J., Ji, F., Klitgord, N., Cusick, M.E., Radivojac, P., Uversky, V.N., Vidal, M., and Iakoucheva, L.M. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput. Biol. *2*, 0890–0901.

Heaton, L.A., Lee, T.C., Wei, N., and Morris, T.J. (1991). Point mutations in the turnip crinkle virus capsid protein affect the symptoms expressed by Nicotiana benthamiana. Virology *183*, 143–150.

Hébrard, E., Bessin, Y., Michon, T., Longhi, S., Uversky, V.N., Delalande, F., Van Dorsselaer, A., Romero, P., Walter, J., Declerk, N., et al. (2009). Intrinsic disorder in Viral Proteins Genome-Linked: experimental and predictive analyses. Virol. J. *6*.

Hébrard, E., Poulicard, N., Gérard, C., Traoré, O., Wu, H.C., Albar, L., Fargette, D., Bessin, Y., and Vignols, F. (2010). Direct interaction between the Rice yellow mottle virus (RYMV) VPg and the central domain of the rice eIF(iso)4G1 factor correlates with rice susceptibility and RYMV virulence. Mol. Plant. Microbe. Interact. *23*, 1506–1513.

Hématy, K., Bellec, Y., Podicheti, R., Bouteiller, N., Anne, P., Morineau, C., Haslam, R.P., Beaudoin, F., Napier, J.A., Mockaitis, K., et al. (2016). The Zinc-Finger Protein SOP1 Is Required for a Subset of the Nuclear Exosome Functions in Arabidopsis. PLoS Genet. *12*.

Heo, I., Ha, M., Lim, J., Yoon, M.-J., Park, J.-E., Kwon, S.C., Chang, H., and Kim, V.N. (2012). Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. Cell *151*, 521–532.

Herzog, E., Guilley, H., and Fritsch, C. (1995). Translation of the Second Gene of Peanut Clump Virus RNA 2 Occurs by Leaky Scanning in Vitro. Virology *208*, 215–225.

Higashimoto, K., Maeda, T., Okada, J., Ohtsuka, Y., Sasaki, K., Hirose, A., Nomiyama, M., Takayanagi, T., Fukuzawa, R., Yatsuki, H., et al. (2013). Homozygous deletion of DIS3L2 exon 9 due to non-allelic homologous recombination between LINE-1s in a Japanese patient with Perlman syndrome. Eur. J. Hum. Genet. *21*, 1316–1319.

Hilcenko, C., Simpson, P.J., Finch, A.J., Bowler, F.R., Churcher, M.J., Jin, L., Packman, L.C., Shlien, A., Campbell, P., Kirwan, M., et al. (2013). Aberrant 3' oligoadenylation of spliceosomal U6 small nuclear RNA in poikiloderma with neutropenia. Blood *121*, 1028–1038.

Hill, K.R., Hajjou, M., Hu, J.Y., and Raju, R. (1997). RNA-RNA recombination in Sindbis virus: roles of the 3' conserved motif, poly(A) tail, and nonviral sequences of template RNAs in polymerase recognition and template switching. J. Virol. *71*, 2693–2704.

Hilleren, P.J., and Parker, R. (2003). Cytoplasmic degradation of splice-defective pre-mRNAs and intermediates. Mol. Cell *12*,

1453–1465.

Hillman, B.I., Carrington, J.C., and Morris, T.J. (1987). A defective interfering RNA that contains a mosaic of a plant virus genome. Cell *51*, 427–433.

Himber, C., Dunoyer, P., Moissiard, G., Ritzenthaler, C., and Voinnet, O. (2003). Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. EMBO J. *22*, 4523.

Hirayama, T. (2014). A unique system for regulating mitochondrial mRNA poly(A) status and stability in plants. Plant Signal. Behav. *9*, 1–4.

Hirayama, T., Matsuura, T., Ushiyama, S., Narusaka, M., Kurihara, Y., Yasuda, M., Ohtani, M., Seki, M., Demura, T., Nakashita, H., et al. (2013). A poly(A)-specific ribonuclease directly regulates the poly(A) status of mitochondrial mRNA in Arabidopsis. Nat. Commun. *4*.

Hizi, A., Henderson, L.E., Copeland, T.D., Sowder, R.C., Hixson, C. V., and Oroszlan, S. (1987). Characterization of mouse mammary tumor virus gag-pro gene products and the ribosomal frameshift site by protein sequencing. Proc. Natl. Acad. Sci. U. S. A. *84*, 7041–7045.

Hoefig, K.P., Rath, N., Heinz, G.A., Wolf, C., Dameris, J., Schepers, A., Kremmer, E., Ansel, K.M., and Heissmeyer, V. (2013). Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. Nat. Struct. Mol. Biol. *20*, 73–81.

Hogenhout, S.A., Ammar, E.D., Whitfield, A.E., and Redinbaugh, M.G. (2008). Insect Vector Interactions with Persistently Transmitted Viruses*. Http://Dx.Doi.Org/10.1146/Annurev.Phyto.022508.092135 *46*, 327–359.

Hogle, J.M., Maeda, A., and Harrison, S.C. (1986). Structure and assembly of turnip crinkle virus: I. X-ray crystallographic structure analysis at 3.2 Å resolution. J. Mol. Biol. *191*, 625–638.

Holmes, F.O. (1946). A comparison of the experimental host ranges of Tobacco-etch and Tobacco-mosaic viruses. Phytopathology *36*, 643–659.

Van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science *295*, 2262–2264.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol. *17*, 1265–1272.

Hoshino, S.I., Imai, M., Kobayashi, T., Uchida, N., and Katada, T. (1999). The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3'-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. J. Biol. Chem. *274*, 16677–16680.

Hou, Y.M. (2000). Unusual synthesis by the Escherichia coli CCA-adding enzyme. RNA *6*, 1031–1043.

Hsu, C.L., and Stevens, A. (1993). Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. Mol. Cell. Biol. *13*, 4826–4835.

Hu, S.F., Wei, W.L., Hong, S.F., Fang, R.Y., Wu, H.Y., Lin, P.C., Sanobar, N., Wang, H.P., Sulistio, M., Wu, C.T., et al. (2020). Investigation of the effects of P1 on HC-pro-mediated gene silencing suppression through genetics and omics approaches. Bot. Stud. *61*.

Huang, T.S., Wei, T., Laliberté, J.F., and Wang, A. (2010). A host RNA helicase-like protein, AtRH8, interacts with the potyviral genome-linked protein, VPg, associates with the virus accumulation complex, and is essential for infection. Plant Physiol. *152*, 255–266.

Hung, C.J., Hu, C.C., Lin, N.S., Lee, Y.C., Meng, M., Tsai, C.H., and Hsu, Y.H. (2014). Two key arginine residues in the coat protein of Bamboo mosaic virus differentially affect the accumulation of viral genomic and subgenomic RNAs. Mol. Plant Pathol. *15*, 196–210.

Hunt, A.G. (2020). mRNA 3' end formation in plants: Novel connections to growth, development and environmental responses. Wiley Interdiscip. Rev. RNA *11*.

Hunt, A.G., Xu, R., Addepalli, B., Rao, S., Forbes, K.P., Meeks, L.R., Xing, D., Mo, M., Zhao, H., Bandyopadhyay, A., et al. (2008). Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. BMC Genomics *9*.

Hunt, A.G., Xing, D., and Li, Q.Q. (2012). Plant polyadenylation factors: conservation and variety in the polyadenylation complex in plants. BMC Genomics *13*.

Huo, Y., Shen, J., Wu, H., Zhang, C., Guo, L., Yang, J., and Li, W. (2016). Widespread 3'-end uridylation in eukaryotic RNA viruses. Sci. Rep. *6*, 25454.

Hwang, Y.T., McCartney, A.W., Gidda, S.K., and Mullen, R.T. (2008). Localization of the Carnation Italian ringspot virus replication protein p36 to the mitochondrial outer membrane is mediated by an internal targeting signal and the TOM complex. BMC Cell Biol. *9*.

Hwang, Y.T., Kalischuk, M., Fusaro, A.F., Waterhouse, P.M., and Kawchuk, L. (2013). Small RNA sequencing of Potato leafroll virus-infected plants reveals an additional subgenomic RNA encoding a sequence-specific RNA-binding protein. Virology *438*, 61–69.

Hyodo, K., and Okuno, T. (2014). Host factors used by positive-strand RNA plant viruses for genome replication. J. Gen. Plant Pathol. *80*, 123–135.

Ibrahim, F., Rohr, J., Jeong, W.J., Hesson, J., and Cerutti, H. (2006). Untemplated oligoadenylation promotes degradation of

RISC-cleaved transcripts. Science *314*, 1893.

Ibrahim, F., Rymarquis, L.A., Kim, E.J., Becker, J., Balassa, E., Green, P.J., and Cerutti, H. (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. Proc. Natl. Acad. Sci. U. S. A. *107*, 3906–3911.

ICTV (2022). ICTV.

Ilyas, M., Du, Z., and Simon, A.E. (2021). Opium Poppy Mosaic Virus Has an Xrn-Resistant, Translated Subgenomic RNA and a BTE 3' CITE. J. Virol. *95*.

Inoue-Nagata, A.K., Jordan, R., Kreuze, J., Li, F., López-Moya, J.J., Mäkinen, K., Ohshima, K., Wylie, S.J., and Ictv Report Consortium (2022). ICTV Virus Taxonomy Profile: Potyviridae 2022. J. Gen. Virol. *103*.

Inoue, K., Ohno, M., Sakamoto, H., and Shimura, Y. (1989). Effect of the cap structure on pre-mRNA splicing in Xenopus oocyte nuclei. Genes Dev. *3*, 1472–1479.

Ivanowski, D. (1968). Concerning the mosaic of the tobacco plant. In Phytopathological Classics, pp. 27–30.

Iwakawa, H., Mizumoto, H., Nagano, H., Imoto, Y., Takigawa, K., Sarawaneeyaruk, S., Kaido, M., Mise, K., and Okuno, T. (2008). A viral noncoding RNA generated by cis-element-mediated protection against 5'->3' RNA decay represses both cap-independent and cap-dependent translation. J. Virol. *82*, 10162–10174.

Iwasaki, S., Takeda, A., Motose, H., and Watanabe, Y. (2007). Characterization of Arabidopsis decapping proteins AtDCP1 and AtDCP2, which are essential for post-embryonic development. FEBS Lett. *581*, 2455–2459.

Jaag, H.M., and Nagy, P.D. (2009). Silencing of Nicotiana benthamiana Xrn4p exoribonuclease promotes tombusvirus RNA accumulation and recombination. Virology *386*, 344–352.

Jadão, A.S., Krause-Sakate, R., Liberti, D., Pavan, M.A., Echer, M.M., Svanella-Dumas, L., Zerbini, F.M., Candresse, T., and Le Gall, O. (2007). Further characterization of two sequiviruses infecting lettuce and development of specific RT-PCR primers. Arch. Virol. *152*, 999–1007.

Jae Eun, K., and Wickens, M. (2007). A family of poly(U) polymerases. RNA *13*, 860.

Januszyk, K., Liu, Q., and Lima, C.D. (2011). Activities of human RRP6 and structure of the human RRP6 catalytic domain. RNA *17*, 1566–1577.

Jia, J., Lu, W., Liu, B., Fang, H., Yu, Y., Mo, W., Zhang, H., Jin, X., Shu, Y., Long, Y., et al. (2022). An atlas of plant full-length RNA reveals tissue-specific and monocots-dicots conserved regulation of poly(A) tail length. Nat. Plants.

Jia, X., Yuan, S., Wang, Y., Fu, Y., Ge, Y., Ge, Y., Lan, X., Feng, Y., Qiu, F., Li, P., et al. (2017). The role of alternative polyadenylation in the antiviral innate immune response. Nat. Commun. *8*.

Jiang, J., and Laliberté, J.F. (2011). The genome-linked protein VPg of plant viruses-a protein with many partners. Curr. Opin. Virol. *1*, 347–354.

Jiang, S., Jiang, L., Yang, J., Peng, J., Lu, Y., Zheng, H., Lin, L., Chen, J., and Yan, F. (2018). Over-expression of Oryza sativa Xrn4 confers plant resistance to virus infection. Gene *639*, 44–51.

Jiao, X., Doamekpor, S.K., Bird, J.G., Nickels, B.E., Tong, L., Hart, R.P., and Kiledjian, M. (2017). 5' End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. Cell *168*, 1015-1027.e10.

Jin, X., Cao, X., Wang, X., Jiang, J., Wan, J., Laliberté, J.F., and Zhang, Y. (2018). Three-Dimensional Architecture and Biogenesis of Membrane Structures Associated with Plant Virus Replication. Front. Plant Sci. *9*.

Jinek, M., Coyle, S.M., and Doudna, J.A. (2011). Coupled 5′ nucleotide recognition and processivity in Xrn1-mediated mRNA decay. Mol. Cell *41*, 600.

Jones, M., Prasetyaningrum, P., Litthauer, S., Vegliani, F., Wood, M.W., William Battle, M., Dickson, C., and Jones, M.A. (2022). EXORIBONUCLEASE4 integrates metabolic signals induced by osmotic stress into the circadian system. BioRxiv 2022.07.05.498805.

Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A., and Mizgerd, J.P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat. Cell Biol. *11*, 1157–1163.

Jones, M.R., Blahna, M.T., Kozlowski, E., Matsuura, K.Y., Ferrari, J.D., Morris, S.A., Powers, J.T., Daley, G.Q., Quinton, L.J., and Mizgerd, J.P. (2012). Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. PLoS Genet. *8*.

Jungfleisch, J., Chowdhury, A., Alves-Rodrigues, I., Tharun, S., and Díez, J. (2015). The Lsm1-7-Pat1 complex promotes viral RNA translation and replication by differential mechanisms. RNA *21*, 1469–1479.

Jungfleisch, J., Blasco-Moreno, B., and Díez, J. (2016). Use of Cellular Decapping Activators by Positive-Strand RNA Viruses. Viruses *8*.

Jupin, I., Bouzoubaa, S., Richards, K., Jonard, G., and Guilley, H. (1990). Multiplication of beet necrotic yellow vein virus RNA 3 lacking a 3' poly(A) tail is accompanied by reappearance of the poly(A) tail and a novel short U-rich tract preceding it. Virology *178*, 281–284.

Kalisiak, K., Kuliński, T.M., Tomecki, R., Cysewski, D., Pietras, Z., Chlebowski, A., Kowalska, K., and Dziembowski, A. (2017). A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. Nucleic Acids Res. *45*, 2068–2080.

Kamminga, L.M., Luteijn, M.J., Den Broeder, M.J., Redl, S., Kaaij, L.J.T., Roovers, E.F., Ladurner, P., Berezikov, E., and Ketting,

R.F. (2010). Hen1 is required for oocyte development and piRNA stability in zebrafish. EMBO J. *29*, 3688–3700.

Kanazawa, M., Ikeda, Y., Nishihama, R., Yamaoka, S., Lee, N.H., Yamato, K.T., Kohchi, T., and Hirayama, T. (2020). Regulation of the Poly(A) Status of Mitochondrial mRNA by Poly(A)-Specific Ribonuclease Is Conserved among Land Plants. Plant Cell Physiol. *61*, 470–480.

Kao, C.C., and Ahlquist, P. (1992). Identification of the domains required for direct interaction of the helicase-like and polymerase-like RNA replication proteins of brome mosaic virus. J. Virol. *66*, 7293–7302.

Kappel, C., Trost, G., Czesnick, H., Ramming, A., Kolbe, B., Vi, S.L., Bispo, C., Becker, J.D., de Moor, C., and Lenhard, M. (2015). Genome-Wide Analysis of PAPS1-Dependent Polyadenylation Identifies Novel Roles for Functionally Specialized Poly(A) Polymerases in Arabidopsis thaliana. PLoS Genet. *11*.

Käslin, E., and Heyer, W.-D. (1994). A Multifunctional Exonuclease from Vegetative Schizosaccharomyces pombe Cells Exhibiting in Vitro Strand Exchange Activity. J. Biol. Chem. *269*, 14094–14102.

Kastenmayer, J.P., and Green, P.J. (2000). Novel features of the XRN-family in Arabidopsis: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. Proc. Natl. Acad. Sci. U. S. A. *97*, 13985–13990.

Kastenmayer, J.P., Johnson, M.A., and Green, P.J. (2001). Analysis of XRN Orthologs by Complementation of Yeast Mutants and Localization of XRN–GFP Fusion Proteins. Methods Enzymol. *342*, 269–282.

Kaygun, H., and Marzluff, W.F. (2005). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. Nat. Struct. Mol. Biol. *12*, 794–800.

Kempf, B.J., and Barton, D.J. (2015). Picornavirus RNA polyadenylation by 3D(pol), the viral RNA-dependent RNA polymerase. Virus Res. *206*, 3–11.

Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. (2003). Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. EMBO J. *22*, 3705–3714.

Khan, M.A., and Goss, D.J. (2019). Poly (A) binding protein enhances the binding affinity of potyvirus VPg to eukaryotic initiation factor eIF4F and activates in vitro translation. Int. J. Biol. Macromol. *121*, 947–955.

Kim, B., Ha, M., Loeff, L., Chang, H., Simanshu, D.K., Li, S., Fareh, M., Patel, D.J., Joo, C., and Kim, V.N. (2015). TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. EMBO J. *34*, 1801–1815.

Kim, D., Lee, Y. suk, Jung, S.J., Yeo, J., Seo, J.J., Lee, Y.Y., Lim, J., Chang, H., Song, J., Yang, J., et al. (2020a). Viral hijacking of the TENT4-ZCCHC14 complex protects viral RNAs via mixed tailing. Nat. Struct. Mol. Biol. *27*, 581–588.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., and Chang, H. (2020b). The Architecture of SARS-CoV-2 Transcriptome. Cell *181*, 914-921.e10.

Kirino, Y., and Mourelatos, Z. (2007). The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. RNA *13*, 1397–1401.

Kitajima, E.W., and Costa, A.S. (1973). Aggregates of chloroplasts in local lesions induced in Chenopodium quinoa Wild. by turnip mosaic virus. J. Gen. Virol. *20*, 413–416.

Kneller, E.L.P., Rakotondrafara, A.M., and Miller, W.A. (2006). Cap-independent translation of plant viral RNAs. Virus Res. *119*, 63.

Knouf, E.C., Wyman, S.K., and Tewari, M. (2013). The human TUT1 nucleotidyl transferase as a global regulator of microRNA abundance. PLoS One *8*.

Koenig, R., Barends, S., Gultyaev, A.P., Lesemann, D.E., Vetten, H.J., Loss, S., and Pleij, C.W.A. (2005). Nemesia ring necrosis virus: a new tymovirus with a genomic RNA having a histidylatable tobamovirus-like 3' end. J. Gen. Virol. *86*, 1827–1833.

Komínek, P., Glasa, M., and Bryxiová, M. (2005). Analysis of the molecular variability of Grapevine leafroll-associated virus 1 reveals the presence of two distinct virus groups and their mixed occurrence in grapevines. Virus Genes *31*, 247–255.

Konarska, M.M., Padgett, R.A., and Sharp, P.A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. Cell *38*, 731–736.

Kong, J., Wei, M., Li, G., Lei, R., Qiu, Y., Wang, C., Li, Z.H., and Zhu, S. (2018). The cucumber mosaic virus movement protein suppresses PAMP-triggered immune responses in Arabidopsis and tobacco. Biochem. Biophys. Res. Commun. *498*, 395–401.

KØrner, C.J., Klauser, D., Niehl, A., Domínguez-Ferreras, A., Chinchilla, D., Boller, T., Heinlein, M., and Hann, D.R. (2013). The immunity regulator BAK1 contributes to resistance against diverse RNA viruses. Mol. Plant. Microbe. Interact. *26*, 1271–1280.

Kowalinski, E., Kögel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., and Conti, E. (2016). Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. Mol. Cell *63*, 125–134.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. Gene *299*, 1–34.

Krause, M., Niazi, A.M., Labun, K., Torres Cleuren, Y.N., Müller, F.S., and Valen, E. (2019). tailfindr: alignment-free poly(A) length measurement for Oxford Nanopore RNA and DNA sequencing. RNA *25*, 1229–1241.

Kreuze, J.F., Vaira, A.M., Menzel, W., Candresse, T., Zavriev, S.K., Hammond, J., and Ryu, K.H. (2020). ICTV Virus Taxonomy Profile: Alphaflexiviridae. J. Gen. Virol. *101*, 699–700.

Kumakura, N., Otsuki, H., Tsuzuki, M., Takeda, A., and Watanabe, Y. (2013). Arabidopsis AtRRP44A is the functional homolog of Rrp44/Dis3, an exosome component, is essential for viability and is required for RNA processing and degradation. PLoS One *8*.

Kurosaki, T., Miyoshi, K., Myers, J.R., and Maquat, L.E. (2018). NMD-degradome sequencing reveals ribosome-bound intermediates with 3'-end non-templated nucleotides. Nat. Struct. Mol. Biol. *25*, 940–950.

Kurth, H.M., and Mochizuki, K. (2009). 2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in Tetrahymena. RNA *15*, 675–685.

Kushner, D.B., Lindenbach, B.D., Grdzelishvili, V.Z., Noueiry, A.O., Paul, S.M., and Ahlquist, P. (2003). Systematic, genome-wide identification of host genes affecting replication of a positive-strand RNA virus. Proc. Natl. Acad. Sci. U. S. A. *100*, 15764–15769.

Kusov, Y.Y., Gosert, R., and Gauss-Müller, V. (2005). Replication and in vivo repair of the hepatitis A virus genome lacking the poly(A) tail. J. Gen. Virol. *86*, 1363–1368.

Łabno, A., Warkocki, Z., Kuliński, T., Krawczyk, P.S., Bijata, K., Tomecki, R., and Dziembowski, A. (2016). Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. Nucleic Acids Res. *44*, gkw649.

Lackey, P.E., Welch, J.D., and Marzluff, W.F. (2016). TUT7 catalyzes the uridylation of the 3' end for rapid degradation of histone mRNA. RNA *22*, 1673–1688.

Lafleche, D., Bove, C., Dupont, G., Mouches, C., Astier, T., Garnier, M., and Bove, J.-M. (1972). Site of viral RNA replication in the cells of higher plants: TYMV RNA synthesis on the chloroplast outer membrane system. Proc. Fed. Eur. Biochem. Soc. *27*, 43–71.

Lai, W.S., Kennington, E.A., and Blackshear, P.J. (2003). Tristetraprolin and its family members can promote the cell-free deadenylation of AU-rich element-containing mRNAs by poly(A) ribonuclease. Mol. Cell. Biol. *23*, 3798–3812.

Lakatos, L., Szittya, G., Silhavy, D., and Burgyán, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. EMBO J. *23*, 876.

Lange, H., and Gagliardi, D. (2022). Catalytic activities, molecular connections, and biological functions of plant RNA exosome complexes. Plant Cell *34*, 967–988.

Lange, H., Holec, S., Cognat, V., Pieuchot, L., Le Ret, M., Canaday, J., and Gagliardi, D. (2008). Degradation of a Polyadenylated rRNA Maturation By-Product Involves One of the Three RRP6-Like Proteins in Arabidopsis thaliana . Mol. Cell. Biol. *28*, 3038–3044.

Lange, H., Sement, F.M., Canaday, J., and Gagliardi, D. (2009). Polyadenylation-assisted RNA degradation processes in plants. Trends Plant Sci. *14*, 497–504.

Lange, H., Sement, F.M., and Gagliardi, D. (2011). MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in Arabidopsis thaliana. Plant J. *68*, 51–63.

Lange, H., Zuber, H., Sement, F.M., Chicher, J., Kuhn, L., Hammann, P., Brunaud, V., Bérard, C., Bouteiller, N., Balzergue, S., et al. (2014). The RNA helicases AtMTR4 and HEN2 target specific subsets of nuclear transcripts for degradation by the nuclear exosome in Arabidopsis thaliana. PLoS Genet. *10*.

Lange, H., Ndecky, S.Y.A., Gomez-Diaz, C., Pflieger, D., Butel, N., Zumsteg, J., Kuhn, L., Piermaria, C., Chicher, J., Christie, M., et al. (2019). RST1 and RIPR connect the cytosolic RNA exosome to the Ski complex in Arabidopsis. Nat. Commun. *10*.

Lapointe, C.P., and Wickens, M. (2013). The nucleic acid-binding domain and translational repression activity of a Xenopus terminal uridylyl transferase. J. Biol. Chem. *288*, 20723–20733.

Le, M.-T., Kasprzak, W.K., Kim, T., Gao, F., Young, M.Y., Yuan, X., Shapiro, B.A., Seog, J., and Simon, A.E. (2017). Folding behavior of a T-shaped, ribosome-binding translation enhancer implicated in a wide-spread conformational switch. Elife *6*.

Lebreton, A., Tomecki, R., Dziembowski, A., and Séraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nat. 2008 4567224 *456*, 993–996.

Lee, C.C., Lin, T.L., Lin, J.W., Han, Y.T., Huang, Y.T., Hsu, Y.H., and Meng, M. (2015). Promotion of Bamboo Mosaic Virus Accumulation in Nicotiana benthamiana by 5′→3′ Exonuclease NbXRN4. Front. Microbiol. *6*.

Lee, Y.F., Nomoto, A., Detjen, B.M., and Wimmer, E. (1977). A protein covalently linked to poliovirus genome RNA. Proc. Natl. Acad. Sci. U. S. A. *74*, 59–63.

van Leeuwen, H.C., Liefhebber, J.M.P., and Spaan, W.J.M. (2006). Repair and polyadenylation of a naturally occurring hepatitis C virus 3' nontranslated region-shorter variant in selectable replicon cell lines. J. Virol. *80*, 4336–4343.

Legnini, I., Alles, J., Karaiskos, N., Ayoub, S., and Rajewsky, N. (2019). FLAM-seq: full-length mRNA sequencing reveals principles of poly(A) tail length control. Nat. Methods *16*, 879–886.

Leiser, R.M., Ziegler-Graff, V., Reutenauer, A., Herrbach, E., Lemaire, O., Guilley, H., Richards, K., and Jonard, G. (1992). Agroinfection as an alternative to insects for infecting plants with beet western yellows luteovirus. Proc. Natl. Acad. Sci. U. S. A. *89*, 9136–9140.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol. Cell *12*, 675–687.

Lellis, A.D., Kasschau, K.D., Whitham, S.A., and Carrington, J.C. (2002). Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. Curr. Biol. *12*, 1046–1051.

Léonard, S., Chisholm, J., Laliberté, J.F., and Sanfaçon, H. (2002). Interaction in vitro between the proteinase of Tomato ringspot virus (genus Nepovirus) and the eukaryotic translation initiation factor iso4E from Arabidopsis thaliana. J. Gen. Virol. *83*, 2085–2089.

Léonard, S., Viel, C., Beauchemin, C., Daigneault, N., Fortin, M.G., and Laliberte, J.F. (2004). Interaction of VPg-Pro of turnip

mosaic virus with the translation initiation factor 4E and the poly(A)-binding protein in planta. J. Gen. Virol. *85*, 1055–1063.

Li, F., and Wang, A. (2018). RNA decay is an antiviral defense in plants that is counteracted by viral RNA silencing suppressors. PLoS Pathog. *14*.

Li, Y., and Maine, E.M. (2018). The balance of poly(U) polymerase activity ensures germline identity, survival and development in Caenorhabditis elegans. Development *145*.

Li, C.H., Irmer, H., Gudjonsdottir-Planck, D., Freese, S., Salm, H., Haile, S., Estévez, A.M., and Clayton, C. (2006). Roles of a Trypanosoma brucei 5'->3' exoribonuclease homolog in mRNA degradation. RNA *12*, 2171–2186.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. Curr. Biol. *15*, 1501–1507.

Li, R., Yue, J., Zhang, Y., Zhou, L., Hao, W., Yuan, J., Qiang, B., Ding, J.M., Peng, X., and Cao, J.M. (2008). CLOCK/BMAL1 regulates human nocturnin transcription through binding to the E-box of nocturnin promoter. Mol. Cell. Biochem. *317*, 169–177.

Li, W., Zhang, Y., Zhang, C., Pei, X., Wang, Z., and Jia, S. (2014). Presence of poly(A) and poly(A)-rich tails in a positive-strand RNA virus known to lack 3′ poly(A) tails. Virology *454–455*, 1–10.

Li, W. zhe, Qu, F., and Morris, T.J. (1998). Cell-to-cell movement of turnip crinkle virus is controlled by two small open reading frames that function in trans. Virology *244*, 405–416.

Li, X.H., Heaton, L.A., Morris, T.J., and Simon, A.E. (1989). Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated de novo. Proc. Natl. Acad. Sci. U. S. A. *86*, 9173–9177.

Li, X.H., Valdez, P., Olvera, R.E., and Carrington, J.C. (1997). Functions of the tobacco etch virus RNA polymerase (NIb): subcellular transport and protein-protein interaction with VPg/proteinase (NIa). J. Virol. *71*, 1598–1607.

Li, Y.-I., Chen, Y.-J., Hsu, Y.-H., and Meng, M. (2001). Characterization of the AdoMet-dependent guanylyltransferase activity that is associated with the N terminus of bamboo mosaic virus replicase. J. Virol. *75*, 782–788.

Licht, K., Medenbach, J., Lührmann, R., Kambach, C., and Bindereif, A. (2008). 3'-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. RNA *14*, 1532–1538.

Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N. (2014). Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. Cell *159*, 1365–1376.

Lim, S.L., Qu, Z.P., Kortschak, R.D., Lawrence, D.M., Geoghegan, J., Hempfling, A.L., Bergmann, M., Goodnow, C.C., Ormandy, C.J., Wong, L., et al. (2015). HENMT1 and piRNA Stability Are Required for Adult Male Germ Cell Transposon Repression and to Define the Spermatogenic Program in the Mouse. PLoS Genet. *11*.

Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.-H., Yee, B.A., Yeo, G.W., Coller, J., and Pasquinelli, A.E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. Nat. Struct. Mol. Biol. *24*, 1057–1063.

Lin, C.J., Wen, J., Bejarano, F., Hu, F., Bortolamiol-Becet, D., Kan, L., Sanfilippo, P., Kondo, S., and Lai, E.C. (2017). Characterization of a TUTase/RNase complex required for Drosophila gametogenesis. RNA *23*, 284–296.

Lin, M. Der, Jiao, X., Grima, D., Newbury, S.F., Kiledjian, M., and Chou, T. Bin (2008). Drosophila processing bodies in oogenesis. Dev. Biol. *322*, 276–288.

Lisitsky, I., Klaff, P., and Schuster, G. (1996). Addition of destabilizing poly (A)-rich sequences to endonuclease cleavage sites during the degradation of chloroplast mRNA. Proc. Natl. Acad. Sci. U. S. A. *93*, 13398–13403.

Litvak, S., Tarragó, A., Tarragó-Litvak, L., and Allende, J.E. (1973). Elongation factor-viral genome interaction dependent on the aminoacylation of TYMV and TMV RNAs. Nat. New Biol. *241*, 88–90.

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. Science *327*, 94–97.

Liu, M., Xu, R., Merrill, C., Hong, L., Von Lanken, C., Hunt, A.G., and Li, Q.Q. (2014). Integration of developmental and environmental signals via a polyadenylation factor in Arabidopsis. PLoS One *9*.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell *127*, 1223–1237.

Liu, Y., Nie, H., Liu, H., and Lu, F. (2019). Poly(A) inclusive RNA isoform sequencing (PAIso-seq) reveals wide-spread non-adenosine residues within RNA poly(A) tails. Nat. Commun. *10*.

Liudkovska, V., and Dziembowski, A. (2021). Functions and mechanisms of RNA tailing by metazoan terminal nucleotidyltransferases. Wiley Interdiscip. Rev. RNA *12*.

Llave, C. (2010). Virus-derived small interfering RNAs at the core of plant-virus interactions. Trends Plant Sci. *15*, 701–707.

Lobel, J.H., Tibble, R.W., and Gross, J.D. (2019). Pat1 activates late steps in mRNA decay by multiple mechanisms. Proc. Natl. Acad. Sci. U. S. A. *116*, 23512–23517.

Loesch-Fries, L.S., and Hall, T.C. (1982). In vivo aminoacylation of brome mosaic and barley stripe mosaic virus RNAs. Nat. 1982 2985876 *298*, 771–773.

Long, Y., Jia, J., Mo, W., Jin, X., and Zhai, J. (2021). FLEP-seq: simultaneous detection of RNA polymerase II position, splicing status, polyadenylation site and poly(A) tail length at genome-wide scale by single-molecule nascent RNA sequencing. Nat. Protoc. *16*, 4355–4381.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S.,

Dziembowski, A., and Jensen, T.H. (2011). Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. Mol. Cell *43*, 624–637.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. EMBO J. *32*, 1855–1868.

Lubas, M., Andersen, P.R., Schein, A., Dziembowski, A., Kudla, G., and Jensen, T.H. (2015). The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. Cell Rep. *10*, 178–192.

Lück, R., Steger, G., and Riesner, D. (1996). Thermodynamic Prediction of Conserved Secondary Structure: Application to the RRE Element of HIV, the tRNA-like Element of CMV and the mRNA of Prion Protein. J. Mol. Biol. *258*, 813–826.

Lykke-Andersen, J. (2002). Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. Mol. Cell. Biol. *22*, 8114–8121.

Ma, X., Nicole, M.C., Meteignier, L.V., Hong, N., Wang, G., and Moffett, P. (2015). Different roles for RNA silencing and RNA processing components in virus recovery and virus-induced gene silencing in plants. J. Exp. Bot. *66*, 919–932.

Ma, X., Zhou, Y., and Moffett, P. (2019). Alterations in cellular RNA decapping dynamics affect tomato spotted wilt virus cap snatching and infection in Arabidopsis. New Phytol. *224*, 789–803.

Magden, J., Takeda, N., Li, T., Auvinen, P., Ahola, T., Miyamura, T., Merits, A., and Kääriäinen, L. (2001). Virus-specific mRNA capping enzyme encoded by hepatitis E virus. J. Virol. *75*, 6249–6255.

Maier, K.C., Gressel, S., Cramer, P., and Schwalb, B. (2020). Native molecule sequencing by nano-ID reveals synthesis and stability of RNA isoforms. Genome Res. *30*, 1332–1344.

Makinen, K., Makelainen, K., Arshava, N., Tamm, T., Merits, A., Truve, E., Zavriev, S., and Saarma, M. (2000). Characterization of VPg and the polyprotein processing of cocksfoot mottle virus (genus Sobemovirus). J. Gen. Virol. *81*, 2783–2789.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J. *32*, 1842–1854.

Malpica-López, N., Rajeswaran, R., Beknazariants, D., Seguin, J., Golyaev, V., Farinelli, L., and Pooggin, M.M. (2018). Revisiting the Roles of Tobamovirus Replicase Complex Proteins in Viral Replication and Silencing Suppression. Mol. Plant. Microbe. Interact. *31*, 125–144.

Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol. *4*, 223.

Margis, R., Ritzenthaler, C., Reinbolt, J., Pinck, M., and Pinck, L. (1993). Genome organization of grapevine fanleaf nepovirus RNA2 deduced from the 122K polyprotein P2 in vitro cleavage products. J. Gen. Virol. *74*, 1919–1926.

Margis, R., Viry, M., Pinck, M., Bardonnet, N., and Pinck, L. (1994). Differential proteolytic activities of precursor and mature forms of the 24K proteinase of grapevine fanleaf nepovirus. Virology *200*, 79–86.

Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases. RNA *13*, 1834–1849.

Martin, M.T., Cervera, M.T., and Garcia, J.A. (1995). Properties of the active plum pox potyvirus RNA polymerase complex in defined glycerol gradient fractions. Virus Res. *37*, 127–137.

Martînez, J., Ren, Y.G., Nilsson, P., Ehrenberg, M., and Virtanen, A. (2001). The mRNA Cap Structure Stimulates Rate of Poly(A) Removal and Amplifies Processivity of Degradation *. J. Biol. Chem. *276*, 27923–27929.

Marzluff, W.F., and Koreski, K.P. (2017). Birth and Death of Histone mRNAs. Trends Genet. *33*, 745–759.

Mas, A., Alves-Rodrigues, I., Noueiry, A., Ahlquist, P., and Díez, J. (2006). Host deadenylation-dependent mRNA decapping factors are required for a key step in brome mosaic virus RNA replication. J. Virol. *80*, 246–251.

Más, P., and Beachy, R.N. (1999). Replication of tobacco mosaic virus on endoplasmic reticulum and role of the cytoskeleton and virus movement protein in intracellular distribution of viral RNA. J. Cell Biol. *147*, 945–958.

Matsuda, D., and Dreher, T.W. (2004). The tRNA-like structure of Turnip yellow mosaic virus RNA is a 3'-translational enhancer. Virology *321*, 36–46.

Matsuda, D., and Dreher, T.W. (2007). Cap- and initiator tRNA-dependent initiation of TYMV polyprotein synthesis by ribosomes: evaluation of the Trojan horse model for TYMV RNA translation. RNA *13*, 129–137.

Matsuda, D., Dunoyer, P., Hemmer, O., Fritsch, C., and Dreher, T.W. (2000). The valine anticodon and valylatability of Peanut clump virus RNAs are not essential but provide a modest competitive advantage in plants. J. Virol. *74*, 8720–8725.

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.J., Chen, Q., et al. (2017). Reversible methylation of m 6 A m in the 5' cap controls mRNA stability. Nature *541*, 371–375.

May, J.P., Yuan, X., Sawicki, E., and Simon, A.E. (2018). RNA virus evasion of nonsense-mediated decay. PLoS Pathog. *14*.

Mayer, A. (1968). Concerning the mosaic disease of tobacco. In Phytopathological Classics, J. Johnson, ed. pp. 11–24.

Mayhew, D.E., and Ford, R.E. (1974). Detection of ribonuclease-resistant RNA in chloroplasts of corn leaf tissue infected with maize dwarf mosaic virus. Virology *57*, 503–509.

McCormack, J.C., and Simon, A.E. (2004). Biased Hypermutagenesis Associated with Mutations in an Untranslated Hairpin of an RNA Virus. J. Virol. *78*, 7813–7817.

McCormack, J.C., Yuan, X., Yingling, Y.G., Kasprzak, W., Zamora, R.E., Shapiro, B.A., and Simon, A.E. (2008). Structural domains within the 3' untranslated region of Turnip crinkle virus. J. Virol. *82*, 8706–8720.

Meeks, L.R., Addepalli, B., and Hunt, A.G. (2009). Characterization of genes encoding poly(A) polymerases in plants: evidence for duplication and functional specialization. PLoS One *4*.

Meng, B., Martelli, G.P., Golino, D.A., and Fuchs, M. (2017). Grapevine viruses: Molecular biology, diagnostics and management. Grapevine Viruses Mol. Biol. Diagnostics Manag. 1–698.

Menzel, W., and Vetten, H.J. (2008). Complete nucleotide sequence of an isolate of the Anthriscus strain of Parsnip yellow fleck virus. Arch. Virol. *153*, 2173–2175.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., et al. (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Mol. Cell *64*, 520–533.

Merits, A., Kettunen, R., Mäkinen, K., Lampio, A., Auvinen, P., Kääriäinen, L., and Ahola, T. (1999). Virus-specific capping of tobacco mosaic virus RNA: methylation of GTP prior to formation of covalent complex p126-m7GMP. FEBS Lett. *455*, 45–48.

Merret, R., Nagarajan, V.K., Carpentier, M.C., Park, S., Favory, J.J., Descombin, J., Picart, C., Charng, Y.Y., Green, P.J., Deragon, J.M., et al. (2015). Heat-induced ribosome pausing triggers mRNA co-translational decay in Arabidopsis thaliana. Nucleic Acids Res. *43*, 4121–4132.

Meyer, S., Urbanke, C., and Wahle, E. (2002). Equilibrium studies on the association of the nuclear poly(A) binding protein with poly(A) of different lengths. Biochemistry *41*, 6082–6089.

Michon, T., Estevez, Y., Walter, J., German-Retana, S., and Le Gall, O. (2006). The potyviral virus genome-linked protein VPg forms a ternary complex with the eukaryotic initiation factors eIF4E and eIF4G and reduces eIF4E affinity for a mRNA cap analogue. FEBS J. *273*, 1312–1322.

Miller, W.A., and Giedroc, D.P. (2010). Ribosomal Frameshifting in Decoding Plant Viral RNAs. Recoding Expans. Decod. Rules Enriches Gene Expr. *24*, 193.

Mitchell, P., and Tollervey, D. (2003). An NMD Pathway in Yeast Involving Accelerated Deadenylation and Exosome-Mediated 3′→5′ Degradation. Mol. Cell *11*, 1405–1413.

Mitton-Fry, R.M., DeGregorio, S.J., Wang, J., Steitz, T.A., and Steitz, J.A. (2010). Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. Science *330*, 1244–1247.

Mohanty, B.K., and Kushner, S.R. (2000). Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. *97*, 11966–11971.

Molnar, A., Melnyk, C., and Baulcombe, D.C. (2011). Silencing signals in plants: A long journey for small RNAs. Genome Biol. *12*, 1–8.

Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. Curr. Opin. Plant Biol. *15*, 349–357.

Monsion, B., Incarbone, M., Hleibieh, K., Poignavent, V., Ghannam, A., Dunoyer, P., Daeffler, L., Tilsner, J., and Ritzenthaler, C. (2018). Efficient Detection of Long dsRNA in Vitro and in Vivo Using the dsRNA Binding Domain from FHV B2 Protein. Front. Plant Sci. *9*.

Montgomery, T.A., Rim, Y.S., Zhang, C., Dowen, R.H., Phillips, C.M., Fischer, S.E.J., and Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. PLoS Genet. *8*.

Moon, S.L., Anderson, J.R., Kumagai, Y., Wilusz, C.J., Akira, S., Khromykh, A.A., and Wilusz, J. (2012). A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability. RNA *18*, 2029–2040.

Morgan, M., Much, C., DiGiacomo, M., Azzi, C., Ivanova, I., Vitsios, D.M., Pistolic, J., Collier, P., Moreira, P.N., Benes, V., et al. (2017). mRNA 3' uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome. Nature *548*, 347–351.

Morgan, M., Kabayama, Y., Much, C., Ivanova, I., Di Giacomo, M., Auchynnikava, T., Monahan, J.M., Vitsios, D.M., Vasiliauskaitė, L., Comazzetto, S., et al. (2019). A programmed wave of uridylation-primed mRNA degradation is essential for meiotic progression and mammalian spermatogenesis. Cell Res. *29*, 221–232.

Morozov, I.Y., Jones, M.G., Razak, A.A., Rigden, D.J., and Caddick, M.X. (2010). CUCU modification of mRNA promotes decapping and transcript degradation in Aspergillus nidulans. Mol. Cell. Biol. *30*, 460–469.

Morozov, I.Y., Jones, M.G., Gould, P.D., Crome, V., Wilson, J.B., Hall, A.J.W., Rigden, D.J., and Caddick, M.X. (2012). mRNA 3' tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. Mol. Cell. Biol. *32*, 2585–2595.

Morris, M.R., Astuti, D., and Maher, E.R. (2013). Perlman syndrome: overgrowth, Wilms tumor predisposition and DIS3L2. Am. J. Med. Genet. C. Semin. Med. Genet. *163C*, 106–113.

Moteki, S., and Price, D. (2002). Functional coupling of capping and transcription of mRNA. Mol. Cell *10*, 599–609.

Movahed, N., Patarroyo, C., Sun, J., Vali, H., Laliberté, J.F., and Zheng, H. (2017). Cylindrical Inclusion Protein of Turnip Mosaic Virus Serves as a Docking Point for the Intercellular Movement of Viral Replication Vesicles. Plant Physiol. *175*, 1732–1744.

Mroczek, S., Krwawicz, J., Kutner, J., Lazniewski, M., Kucinski, I., Ginalski, K., and Dziembowski, A. (2012). C16orf57, a gene mutated in poikiloderma with neutropenia, encodes a putative phosphodiesterase responsible for the U6 snRNA 3' end modification. Genes Dev. *26*, 1911–1925.

Muhlrad, D., Decker, C.J., and Parker, R. (1994). Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. Genes Dev. *8*, 855–866.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and

simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. Genes Dev. *22*, 50–65.

Mulot, M., Monsion, B., Boissinot, S., Rastegar, M., Meyer, S., Bochet, N., and Brault, V. (2018). Transmission of Turnip yellows virus by Myzus persicae Is Reduced by Feeding Aphids on Double-Stranded RNA Targeting the Ephrin Receptor Protein. Front. Microbiol. *9*.

Munoz-Tello, P., Gabus, C., and Thore, S. (2012). Functional implications from the Cid1 poly(U) polymerase crystal structure. Structure *20*, 977–986.

Murphy, J.F., Klein, P.G., Hunt, A.G., and Shaw, J.G. (1996). Replacement of the tyrosine residue that links a potyviral VPg to the viral RNA is lethal. Virology *220*, 535–538.

Na, H., Fabian, M.R., and White, K.A. (2006). Conformational organization of the 3′ untranslated region in the tomato bushy stunt virus genome. RNA *12*, 2199.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'→3' exoribonucleases: structure, mechanisms and functions. Biochim. Biophys. Acta *1829*, 590–603.

Nagarajan, V.K., Kukulich, P.M., Von Hagel, B., and Green, P.J. (2019). RNA degradomes reveal substrates and importance for dark and nitrogen stress responses of Arabidopsis XRN4. Nucleic Acids Res. *47*, 9216–9230.

Nagy, P.D., Carpenter, C.D., and Simon, A.E. (1997). A novel 3'-end repair mechanism in an RNA virus. Proc. Natl. Acad. Sci. U. S. A. *94*, 1113–1118.

Nellist, C.F., Ohshima, K., Ponz, F., and Walsh, J.A. (2022). Turnip mosaic virus, a virus for all seasons. Ann. Appl. Biol. *180*, 312–327.

Neufeld, K.L., Galarza, J.M., Richards, O.C., Summers, D.F., and Ehrenfeld, E. (1994). Identification of terminal adenylyl transferase activity of the poliovirus polymerase 3Dpol. J. Virol. *68*, 5811–5818.

Newbury, S., and Woollard, A. (2004). The 5'-3' exoribonuclease xrn-1 is essential for ventral epithelial enclosure during C. elegans embryogenesis. RNA *10*, 59–65.

Newman, M.A., Mani, V., and Hammond, S.M. (2011). Deep sequencing of microRNA precursors reveals extensive 3' end modification. RNA *17*, 1795–1803.

Ni, P., Vaughan, R.C., Tragesser, B., Hoover, H., and Kao, C.C. (2014). The plant host can affect the encapsidation of brome mosaic virus (BMV) RNA: BMV virions are surprisingly heterogeneous. J. Mol. Biol. *426*, 1061–1076.

Nicaise, V., and Candresse, T. (2017). Plum pox virus capsid protein suppresses plant pathogen-associated molecular pattern (PAMP)-triggered immunity. Mol. Plant Pathol. *18*, 878–886.

Nicholson-Shaw, A.L., Kofman, E.R., Yeo, G.W., and Pasquinelli, A.E. (2022). Nuclear and cytoplasmic poly(A) binding proteins (PABPs) favor distinct transcripts and isoforms. Nucleic Acids Res. *50*, 4685–4702.

Nicolas, O., and Laliberte, J.F. (1992). The complete nucleotide sequence of turnip mosaic potyvirus RNA. J. Gen. Virol. *73*, 2785–2793.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol. Cell *39*, 773–783.

Nojima, T., Hirose, T., Kimura, H., and Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J. Biol. Chem. *282*, 15645–15651.

Noueiry, A.O., Diez, J., Falk, S.P., Chen, J., and Ahlquist, P. (2003). Yeast Lsm1p-7p/Pat1p deadenylation-dependent mRNA-decapping factors are required for brome mosaic virus genomic RNA translation. Mol. Cell. Biol. *23*, 4094–4106.

Nowak, J.S., Hobor, F., Velasco, A.D.R., Choudhury, N.R., Heikel, G., Kerr, A., Ramos, A., and Michlewski, G. (2017). Lin28a uses distinct mechanisms of binding to RNA and affects miRNA levels positively and negatively. RNA *23*, 317–332.

Ochi, H., and Chiba, K. (2016). Hormonal stimulation of starfish oocytes induces partial degradation of the 3' termini of cyclin B mRNAs with oligo(U) tails, followed by poly(A) elongation. RNA *22*, 822–829.

Ogami, K., Richard, P., Chen, Y., Hoque, M., Li, W., Moresco, J.J., Yates, J.R., Tian, B., and Manley, J.L. (2017). An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. Genes Dev. *31*, 1257–1271.

Okade, H., Fujita, Y., Miyamoto, S., Tomoo, K., Muto, S., Miyoshi, H., Natsuaki, T., Rhoads, R.E., and Ishida, T. (2009). Turnip mosaic virus genome-linked protein VPg binds C-terminal region of cap-bound initiation factor 4E orthologue without exhibiting host cellular specificity. J. Biochem. *145*, 299–307.

Olspert, A., Peil, L., Hébrard, E., Fargette, D., and Truve, E. (2011a). Protein-RNA linkage and post-translational modifications of two sobemovirus VPgs. J. Gen. Virol. *92*, 445–452.

Olspert, A., Arike, L., Peil, L., and Truve, E. (2011b). Sobemovirus RNA linked to VPg over a threonine residue. FEBS Lett. *585*, 2979–2985.

Olsthoorn, R. (2022). Replication of alphaviruses requires a pseudoknot that involves the polyA tail. BioRxiv 2022.06.20.496828.

Olsthoorn, R.C.L., Mertens, S., Brederode, F.T., and Bol, J.F. (1999). A conformational switch at the 3' end of a plant virus RNA regulates viral replication. EMBO J. *18*, 4856–4864.

van Ooij, M.J.M., Polacek, C., Glaudemans, D.H.R.F., Kuijpers, J., van Kuppeveld, F.J.M., Andino, R., Agol, V.I., and Melchers, W.J.G. (2006). Polyadenylation of genomic RNA and initiation of antigenomic RNA in a positive-strand RNA virus are controlled by the same cis-element. Nucleic Acids Res. *34*, 2953–2965.

Orban, T.I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA *11*, 459–469.

Pabis, M., Neufeld, N., Steiner, M.C., Bojic, T., Shav-Tal, Y., and Neugebauer, K.M. (2013). The nuclear cap-binding complex interacts with the U4/U6·U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. RNA *19*, 1054–1063.

Palukaitis, P. (2016). Satellite RNAs and Satellite Viruses. Mol. Plant. Microbe. Interact. *29*, 181–186.

Panavas, T., Pogany, J., and Nagy, P.D. (2002). Internal initiation by the cucumber necrosis virus RNA-dependent RNA polymerase is facilitated by promoter-like sequences. Virology *296*, 275–287.

Pandey, N.B., and Marzluff, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol. Cell. Biol. *7*, 4557–4559.

Pantaleo, V., Szittya, G., and Burgyán, J. (2007). Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. J. Virol. *81*, 3797–3806.

Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. Nat. Struct. Mol. Biol. 2004 112 *11*, 121–127.

Parker, M.T., Knop, K., Sherwood, A. V., Schurch, N.J., Mackinnon, K., Gould, P.D., Hall, A.J.W., Barton, G.J., and Simpson, G.G. (2020). Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m 6 A modification. Elife *9*.

Passmore, L.A., and Coller, J. (2021). Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. Nat. Rev. Mol. Cell Biol. 2021 232 *23*, 93–106.

Paul, A. V., and Wimmer, E. (2015). Initiation of protein-primed picornavirus RNA synthesis. Virus Res. *206*, 12–26.

Paul, A. V., Van Boom, J.H., Filippov, D., and Wimmer, E. (1998). Protein-primed RNA synthesis by purified poliovirus RNA polymerase. Nature *393*, 280–284.

Pavlopoulou, A., Vlachakis, D., Balatsos, N.A.A., and Kossida, S. (2013). A comprehensive phylogenetic analysis of deadenylases. Evol. Bioinform. Online *9*, 491–497.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2015). Widespread co-translational RNA decay reveals ribosome dynamics. Cell *161*, 1400–1412.

Pelletier, J., Kaplan, G., Racaniello, V.R., and Sonenberg, N. (1988). Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. Mol. Cell. Biol. *8*, 1103–1112.

Peltier, C., Klein, E., Hleibieh, K., D'Alonzo, M., Hammann, P., Bouzoubaa, S., Ratti, C., and Gilmer, D. (2012). Beet necrotic yellow vein virus subgenomic RNA3 is a cleavage product leading to stable non-coding RNA required for long-distance movement. J. Gen. Virol. *93*, 1093–1102.

Le Pen, J., Jiang, H., Di Domenico, T., Kneuss, E., Kosałka, J., Leung, C., Morgan, M., Much, C., Rudolph, K.L.M., Enright, A.J., et al. (2018). Terminal uridylyltransferases target RNA viruses as part of the innate immune system. Nat. Struct. Mol. Biol. *25*, 778–786.

Peng, J., Yang, J., Yan, F., Lu, Y., Jiang, S., Lin, L., Zheng, H., Chen, H., and Chen, J. (2011). Silencing of NbXrn4 facilitates the systemic infection of Tobacco mosaic virus in Nicotiana benthamiana. Virus Res. *158*, 268–270.

Peng, Y.H., Lin, C.H., Lin, C.N., Lo, C.Y., Tsai, T.L., and Wu, H.Y. (2016). Characterization of the Role of Hexamer AGUAAA and Poly(A) Tail in Coronavirus Polyadenylation. PLoS One *11*, e0165077.

Penn, W.D., Harrington, H.R., Schlebach, J.P., and Mukhopadhyay, S. (2020). Regulators of Viral Frameshifting: More Than RNA Influences Translation Events. Annu. Rev. Virol. *7*, 219–238.

Perea-Resa, C., Hernández-Verdeja, T., López-Cobollo, R., Castellano, M. del M., and Salinas, J. (2013). LSM proteins provide accurate splicing and decay of selected transcripts to ensure normal Arabidopsis development. Plant Cell *24*, 4930–4947.

Pérez-Cañamás, M., Hevia, E., Katsarou, K., and Hernández, C. (2021). Genetic evidence for the involvement of Dicer-like 2 and 4 as well as Argonaute 2 in the Nicotiana benthamiana response against Pelargonium line pattern virus. J. Gen. Virol. *102*.

Perraki, A., Gronnier, J., Gouguet, P., Boudsocq, M., Deroubaix, A.F., Simon, V., German-Retana, S., Legrand, A., Habenstein, B., Zipfel, C., et al. (2018). REM1.3's phospho-status defines its plasma membrane nanodomain organization and activity in restricting PVX cell-to-cell movement. PLoS Pathog. *14*.

Petrillo, J.E., Rocheleau, G., Kelley-Clarke, B., and Gehrke, L. (2005). Evaluation of the conformational switch model for alfalfa mosaic virus RNA replication. J. Virol. *79*, 5743–5751.

Picard-Jean, F., Tremblay-Létourneau, M., Serra, E., Dimech, C., Schulz, H., Anselin, M., Dutilly, V., and Bisaillon, M. (2013). RNA 5′-end Maturation: A Crucial Step in the Replication of Viral Genomes. Curr. Issues Mol. Virol. - Viral Genet. Biotechnol. Appl.

Pijlman, G.P., Funk, A., Kondratieva, N., Leung, J., Torres, S., van der Aa, L., Liu, W.J., Palmenberg, A.C., Shi, P.Y., Hall, R.A., et al. (2008). A Highly Structured, Nuclease-Resistant, Noncoding RNA Produced by Flaviviruses Is Required for Pathogenicity. Cell Host Microbe *4*, 579–591.

Pinck, L., Fuchs, . M, Pinck, M., Ravelonandroi~, M., and Walter, B. (1988). A Satellite RNA in Grapevine Fanleaf Virus Strain F13.

Pinck, M., Yot, P., Chapeville, F., and Duranton, H.M. (1970). Enzymatic binding of valine to the 3' end of TYMV-RNA. Nature *226*, 954–956.

Pinck, M., Reinbolt, J., Loudes, A.M., Le Ret, M., and Pinck, L. (1991). Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. FEBS Lett. *284*, 117–119.

Pirouz, M., Du, P., Munafò, M., and Gregory, R.I. (2016). Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. Cell Rep. *16*, 1861–1873.

Pirouz, M., Munafò, M., Ebrahimi, A.G., Choe, J., and Gregory, R.I. (2019). Exonuclease Requirements for Mammalian Ribosomal RNA Biogenesis and Surveillance. Nat. Struct. Mol. Biol. *26*, 490.

Pirouz, M., Wang, C.H., Liu, Q., Ebrahimi, A.G., Shamsi, F., Tseng, Y.H., and Gregory, R.I. (2020). The Perlman syndrome DIS3L2 exoribonuclease safeguards endoplasmic reticulum-targeted mRNA translation and calcium ion homeostasis. Nat. Commun. *11*.

Pisacane, P., and Halic, M. (2017). Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. Nat. Commun. *8*.

Plante, D., Viel, C., Léonard, S., Tampo, H., Laliberté, J.F., and Fortin, M.G. (2004). Turnip mosaic virus VPg does not disrupt the translation initiation complex but interferes with cap binding. Physiol. Mol. Plant Pathol. *64*, 219–226.

Poon, L.L.M., Pritlove, D.C., Fodor, E., and Brownlee, G.G. (1999). Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. J. Virol. *73*, 3473–3476.

Poranen, M.M., Koivunen, M.R.L., and Bamford, D.H. (2008). Nontemplated terminal nucleotidyltransferase activity of double-stranded RNA bacteriophage phi6 RNA-dependent RNA polymerase. J. Virol. *82*, 9254–9264.

Portnoy, V., Evguenieva-Hackenberg, E., Klein, F., Walter, P., Lorentzen, E., Klug, G., and Schuster, G. (2005). RNA polyadenylation in Archaea: not observed in Haloferax while the exosome polynucleotidylates RNA in Sulfolobus. EMBO Rep. *6*, 1188–1193.

Preti, M., O'Donohue, M.F., Montel-Lehry, N., Bortolin-Cavaillé, M.L., Choesmel, V., and Gleizes, P.E. (2013). Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. Nucleic Acids Res. *41*, 4709–4723.

Puustinen, P., and Mäkinen, K. (2004). Uridylylation of the potyvirus VPg by viral replicase NIb correlates with the nucleotide binding capacity of VPg. J. Biol. Chem. *279*, 38103–38110.

Qu, F., and Morris, T.J. (2000). Cap-independent translational enhancement of turnip crinkle virus genomic and subgenomic RNAs. J. Virol. *74*, 1085–1093.

Qu, F., Ren, T., and Morris, T.J. (2003). The coat protein of turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. J. Virol. *77*, 511–522.

Rajamäki, M.L., and Valkonen, J.P.T. (2009). Control of nuclear and nucleolar localization of nuclear inclusion protein a of picorna-like Potato virus A in Nicotiana species. Plant Cell *21*, 2485–2502.

Rajappa-Titu, L., Suematsu, T., Munoz-Tello, P., Long, M., Demir, Ö., Cheng, K.J., Stagno, J.R., Luecke, H., Amaro, R.E., Aphasizheva, I., et al. (2016). RNA Editing TUTase 1: structural foundation of substrate recognition, complex interactions and drug targeting. Nucleic Acids Res. *44*, 10862–10878.

Raju, R., Hajjou, M., Hill, K.R., Botta, V., and Botta, S. (1999). In vivo addition of poly(A) tail and AU-rich sequences to the 3' terminus of the Sindbis virus RNA genome: a novel 3'-end repair pathway. J. Virol. *73*, 2410–2419.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. *44*, 7511–7526.

Ranjith-Kumar, C.T., Gajewski, J., Gutshall, L., Maley, D., Sarisky, R.T., and Kao, C.C. (2001). Terminal nucleotidyl transferase activity of recombinant Flaviviridae RNA-dependent RNA polymerases: implication for viral RNA synthesis. J. Virol. *75*, 8615–8623.

Rantalainen, K.I., Uversky, V.N., Permi, P., Kalkkinen, N., Dunker, A.K., and Mäkinen, K. (2008). Potato virus A genome-linked protein VPg is an intrinsically disordered molten globule-like protein with a hydrophobic core. Virology *377*, 280–288.

Rantalainen, K.I., Christensen, P.A., Hafrén, A., Otzen, D.E., Kalkkinen, N., and Mäkinen, K. (2009). Interaction of a potyviral VPg with anionic phospholipid vesicles. Virology *395*, 114–120.

Rantalainen, K.I., Eskelin, K., Tompa, P., and Mäkinen, K. (2011). Structural flexibility allows the functional diversity of potyvirus genome-linked protein VPg. J. Virol. *85*, 2449–2457.

Rao, A.L., Dreher, T.W., Marsh, L.E., and Hall, T.C. (1989). Telomeric function of the tRNA-like structure of brome mosaic virus RNA. Proc. Natl. Acad. Sci. U. S. A. *86*, 5335–5339.

Ratti, C., Hleibieh, K., Bianchi, L., Schirmer, A., Autonell, C.R., and Gilmer, D. (2009). Beet soil-borne mosaic virus RNA-3 is replicated and encapsidated in the presence of BNYVV RNA-1 and -2 and allows long distance movement in Beta macrocarpa. Virology *385*, 392–399.

Reimão-Pinto, M.M., Manzenreither, R.A., Burkard, T.R., Sledz, P., Jinek, M., Mechtler, K., and Ameres, S.L. (2016). Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in Drosophila. EMBO J. *35*, 2417–2434.

Reinbold, C., Lacombe, S., Ziegler-Graff, V., Scheidecker, D., Wiss, L., Beuve, M., Caranta, C., and Brault, V. (2013). Closely Related Poleroviruses Depend on Distinct Translation Initiation Factors to Infect Arabidopsis thaliana. Http://Dx.Doi.Org/10.1094/MPMI-07-12-0174-R *26*, 257–265.

Ren, G., Chen, X., and Yu, B. (2012). Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Curr. Biol. *22*, 695–700.

Ren, G., Xie, M., Zhang, S., Vinovskis, C., Chen, X., and Yu, B. (2014). Methylation protects microRNAs from an AGO1-associated activity that uridylates 5' RNA fragments generated by AGO1 cleavage. Proc. Natl. Acad. Sci. U. S. A. *111*, 6365–6370.

Restrepo-Hartwig, M.A., and Ahlquist, P. (1996). Brome mosaic virus helicase- and polymerase-like proteins colocalize on the endoplasmic reticulum at sites of viral RNA synthesis. J. Virol. *70*, 8908–8916.

Reverdatto, S. V., Dutko, J.A., Chekanova, J.A., Hamilton, D.A., and Belostotsky, D.A. (2004). mRNA deadenylation by PARN is essential for embryogenesis in higher plants. RNA *10*, 1200–1214.

Revers, F., and García, J.A. (2015). Molecular biology of potyviruses. Adv. Virus Res. *92*, 101–199.

Riechmann, J., Laín, S., and Garcia, J.A. (1990). Infectious in vitro transcripts from a plum pox potyvirus cDNA clone. Virology *177*, 710–716.

Rietveld, K., Pleij, C.W., and Bosch, L. (1983). Three-dimensional models of the tRNA-like 3' termini of some plant viral RNAs. EMBO J. *2*, 1079–1085.

Rietveld, K., Linschooten, K., Pleij, C.W.A., and Bosch, L. (1984). The three-dimensional folding of the tRNA-like structure of tobacco mosaic virus RNA. A new building principle applied twice. EMBO J. *3*, 2613–2619.

Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. Nat. Struct. Mol. Biol. *16*, 616–623.

Ritzenthaler, C., Viry, M., Pinck, M., Margis, R., Fuchs, M., and Pinck, L. (1991). Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. J. Gen. Virol. *72*, 2357–2365.

Ritzenthaler, C., Pinck, M., and Pinck, L. (1995). Grapevine fanleaf nepovirus P38 putative movement protein is not transiently expressed and is a stable final maturation product in vivo. J. Gen. Virol. *76 (Pt 4)*, 907–915.

Ritzenthaler, C., Laporte, C., Gaire, F., Dunoyer, P., Schmitt, C., Duval, S., Piéquet, A., Loudes, A.M., Rohfritsch, O., Stussi-Garaud, C., et al. (2002). Grapevine fanleaf virus replication occurs on endoplasmic reticulum-derived membranes. J. Virol. *76*, 8808–8819.

Roach, N.P., Sadowski, N., Alessi, A.F., Timp, W., Taylor, J., and Kim, J.K. (2020). The full-length transcriptome of C. Elegans using direct RNA sequencing. Genome Res. *30*, 299–312.

Rochon, D., Lommel, S., Martelli, G.P., Rubino, L., and Russo, M. (2011). ICTV Virus Taxonomy Profile: Tombusviridae 2011.

Rohayem, J., Jäger, K., Robel, I., Scheffler, U., Temme, A., and Rudolph, W. (2006). Characterization of norovirus 3Dpol RNA-dependent RNA polymerase activity and initiation of RNA synthesis. J. Gen. Virol. *87*, 2621–2630.

Roossinck, M.J. (2001). Cucumber mosaic virus, a model for RNA virus evolution. Mol. Plant Pathol. *2*, 59–63.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol. Biol. 2003 531 *53*, 247–259.

Rothberg, P.G., Harris, T.J.R., Nomoto, A., and Wimmer, E. (1978). O4-(5'-uridylyl)tyrosine is the bond between the genome-linked protein and the RNA of poliovirus. Proc. Natl. Acad. Sci. U. S. A. *75*, 4868–4872.

Rothnie, H.M., Chen, G., Fütterer, J., and Hohn, T. (2001). Polyadenylation in Rice Tungro Bacilliform Virus: cis -Acting Signals and Regulation . J. Virol. *75*, 4184–4194.

Rott, R., Zipor, G., Portnoy, V., Liveanu, V., and Schuster, G. (2003). RNA polyadenylation and degradation in cyanobacteria are similar to the chloroplast but different from Escherichia coli. J. Biol. Chem. *278*, 15771–15777.

Roux, M.E., Rasmussen, M.W., Palma, K., Lolle, S., Regué, À.M., Bethke, G., Glazebrook, J., Zhang, W., Sieburth, L., Larsen, M.R., et al. (2015). The mRNA decay factor PAT1 functions in a pathway including MAP kinase 4 and immune receptor SUMM2. EMBO J. *34*, 593–608.

de Rozières, C.M., and Joseph, S. (2020). Influenza A Virus NS1 Protein Binds as a Dimer to RNA-Free PABP1 but Not to the PABP1·Poly(A) RNA Complex. Biochemistry *59*, 4439–4448.

Rudinger-Thirion, J., Olsthoorn, R.C.L., Giegé, R., and Barends, S. (2006). Idiosyncratic behaviour of tRNA-like structures in translation of plant viral RNA genomes. J. Mol. Biol. *355*, 873–878.

Rudinger, J., Felden, B., Florentz, C., and Giegé, R. (1997). Strategy for RNA recognition by yeast histidyl-tRNA synthetase. Bioorg. Med. Chem. *5*, 1001–1009.

Russo, M., and Martelli, G.P. (1982). Ultrastructure of turnip crinkle- and saguaro cactus virus-infected tissues. Virology *118*, 109–116.

Rymarquis, L.A., Souret, F.F., and Green, P.J. (2011). Evidence that XRN4, an Arabidopsis homolog of exoribonuclease XRN1, preferentially impacts transcripts with certain sequences or in particular functional categories. RNA *17*, 501–511.

Sachs, A.B., Bond, M.W., and Kornberg, R.D. (1986). A Single Gene From Yeast for Both Nuclear and Cytoplasmic Polyadenylate-Binding Proteins: Domain Structure and Expression. Cell *45*, 827–835.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev. *21*, 1603–1608.

Sakuno, T., Araki, Y., Ohya, Y., Kofuji, S., Takahashi, S., Hoshino, S.I., and Katada, T. (2004). Decapping reaction of mRNA requires Dcp1 in fission yeast: its characterization in different species from yeast to human. J. Biochem. *136*, 805–812.

Sanjuán, R., and Domingo-Calap, P. (2016). Mechanisms of viral mutation. Cell. Mol. Life Sci. *73*, 4433–4448.

Sastry, K.S., and Zitter, T.A. (2014). Plant virus and viroid diseases in the tropics: Volume 2: Epidemiology and management. Plant Virus Viroid Dis. Trop. Vol. 2 Epidemiol. Manag. 1–489.

Satheshkumar, P.S., Gayathri, P., Prasad, K., and Savithri, H.S. (2005). "Natively unfolded" VPg is essential for Sesbania mosaic

virus serine protease activity. J. Biol. Chem. *280*, 30291–30300.

Sato, H., and Maquat, L.E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin β. Genes Dev. *23*, 2537.

Sato, M., Nakahara, K., Yoshii, M., Ishikawa, M., and Uyeda, I. (2005). Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of Arabidopsis thaliana by potyviruses. FEBS Lett. *579*, 1167–1171.

Sato, Y., Shimamoto, A., Shobuike, T., Sugimoto, M., Ikeda, H., Kuroda, S., and Furuichi, Y. (1998). Cloning and characterization of human Sep1 (hSEP1) gene and cytoplasmic localization of its product. DNA Res. *5*, 241–246.

Sawicki, D.L., and Gomatos, P.J. (1976). Replication of semliki forest virus: polyadenylate in plus-strand RNA and polyuridylate in minus-strand RNA. J. Virol. *20*, 446–464.

Schaad, M.C., Anderberg, R.J., and Carrington, J.C. (2000). Strain-specific interaction of the tobacco etch virus NIa protein with the translation initiation factor eIF4E in the yeast two-hybrid system. Virology *273*, 300–306.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M., and Van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat. Struct. Mol. Biol. *16*, 56–62.

Schäfer, I.B., Yamashita, M., Schuller, J.M., Schüssler, S., Reichelt, P., Strauss, M., and Conti, E. (2019). Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. Cell *177*, 1619-1631.e21.

Scheer, H., De Almeida, C., Sikorska, N., Koechler, S., Gagliardi, D., and Zuber, H. (2020). High-Resolution Mapping of 3' Extremities of RNA Exosome Substrates by 3' RACE-Seq. Methods Mol. Biol. *2062*, 147–167.

Scheer, H., De Almeida, C., Ferrier, E., Simonnot, Q., Poirier, L., Pflieger, D., Sement, F.M., Koechler, S., Piermaria, C., Krawczyk, P., et al. (2021). The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis. Nat. Commun. *12*.

Schiaffini, M., Chicois, C., Pouclet, A., Chartier, T., Ubrig, E., Gobert, A., Zuber, H., Mutterer, J., Chicher, J., Kuhn, L., et al. (2022). A NYN domain protein directly interacts with DECAPPING1 and is required for phyllotactic pattern. Plant Physiol. *188*, 1174–1188.

Schmid, M., and Jensen, T.H. (2019). The Nuclear RNA Exosome and Its Cofactors. Adv. Exp. Med. Biol. *1203*, 113–132.

Schmid, M., Poulsen, M.B., Olszewski, P., Pelechano, V., Saguez, C., Gupta, I., Steinmetz, L.M., Moore, C., and Jensen, T.H. (2012). Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. Mol. Cell *47*, 267–280.

Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., et al. (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. Science *354*, 1431–1433.

Schneider, C., Leung, E., Brown, J., and Tollervey, D. (2009). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res. *37*, 1127–1140.

Scholthof, K.B.G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., et al. (2011). Top 10 plant viruses in molecular plant pathology. Mol. Plant Pathol. *12*, 938–954.

Schultz, E.S. (1921). A transmissible mosaic disease of chinese cabbage, mustard, and turnip. J. Agric. Res. *XXII*.

Sement, F.M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.-M., Bousquet-Antonelli, C., Lange, H., and Gagliardi, D. (2013). Uridylation prevents 3' trimming of oligoadenylated mRNAs. Nucleic Acids Res. *41*, 7115–7127.

Serghini, M.A., Fuchs, M., Pinck, M., Reinbolt, J., Walter, B., and Pinck, L. (1990). RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location. J. Gen. Virol. *71*, 1433–1441.

Seth, M., Thurlow, D.L., and Hou, Y.M. (2002). Poly(C) synthesis by class I and class II CCA-adding enzymes. Biochemistry *41*, 4521–4532.

Sharif, H., and Conti, E. (2013). Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover. Cell Rep. *5*, 283–291.

Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: capping and polyadenylation. Nat. Struct. Biol. *7*, 838–842.

Shchepachev, V., Wischnewski, H., Missiaglia, E., Soneson, C., and Azzalin, C.M. (2012). Mpn1, Mutated in Poikiloderma with Neutropenia Protein 1, Is a Conserved 3′-to-5′ RNA Exonuclease Processing U6 Small Nuclear RNA. Cell Rep. *2*, 855–865.

She, M., Decker, C.J., Sundramurthy, K., Liu, Y., Chen, N., Parker, R., and Song, H. (2004). Crystal structure of Dcp1p and its functional implications in mRNA decapping. Nat. Struct. Mol. Biol. *11*, 249–256.

She, M., Decker, C.J., Svergun, D.I., Round, A., Chen, N., Muhlrad, D., Parker, R., and Song, H. (2008). Structural basis of dcp2 recognition and activation by dcp1. Mol. Cell *29*, 337–349.

Shen, B., and Goodman, H.M. (2004). Uridine addition after microRNA-directed cleavage. Science *306*, 997.

Sherlock, M.E., Hartwick, E.W., MacFadden, A., and Kieft, J.S. (2021). Structural diversity and phylogenetic distribution of valyl tRNA-like structures in viruses. Rna *27*, 27–39.

Shi, H., Barnes, R.L., Carriero, N., Atayde, V.D., Tschudi, C., and Ullu, E. (2014). Role of the Trypanosoma brucei HEN1 family methyltransferase in small interfering RNA modification. Eukaryot. Cell *13*, 77–86.

Shien, J.H., Su, Y. Da, and Wu, H.Y. (2014). Regulation of coronaviral poly(A) tail length during infection is not coronavirus species- or host cell-specific. Virus Genes *49*, 383–392.

Shimizu, K., Iguchi, A., Gomyou, R., and Yasushi, O. (1999). Influenza virus inhibits cleavage of the HSP70 pre-mRNAs at the polyadenylation site. Virology *254*, 213–219.

Shin, J., Paek, K.Y., Chikhaoui, L., Jung, S., Ponny, S.R., Suzuki, Y., Padmanabhan, K., and Richter, J.D. (2022). Oppositional poly(A) tail length regulation by FMRP and CPEB1. RNA *28*, 756–765.

Shobuike, T., Sugano, S., Yamashita, T., and Ikeda, H. (1995). Characterization of cDNA encoding mouse homolog of fission yeast dhp1+ gene: structural and functional conservation. Nucleic Acids Res. *23*, 357–361.

Shobuike, T., Tatebayashi, K., Tani, T., Sugano, S., and Ikeda, H. (2001). The dhp1(+) gene, encoding a putative nuclear 5'-->3' exoribonuclease, is required for proper chromosome segregation in fission yeast. Nucleic Acids Res. *29*, 1326–1333.

Siddiqui, N., Osborne, M.J., Gallie, D.R., and Gehring, K. (2007). Solution structure of the PABC domain from wheat poly (A)-binding protein: An insight into RNA metabolic and translational control in plants. Biochemistry *46*, 4221–4231.

Sikorska, N. (2016). The phosphorolytic activity of the exosome core complex contributes to rRNA maturation in Arabidopsis thaliana. University of Strasbourg.

Sikorska, N., Zuber, H., Gobert, A., Lange, H., and Gagliardi, D. (2017). RNA degradation by the plant RNA exosome involves both phosphorolytic and hydrolytic activities. Nat. Commun. *8*, 2162.

Sikorski, P.J., Zuber, H., Philippe, L., Sement, F.M., Canaday, J., Kufel, J., Gagliardi, D., and Lange, H. (2015). Distinct 18S rRNA precursors are targets of the exosome complex, the exoribonuclease RRP6L2 and the terminal nucleotidyltransferase TRL in Arabidopsis thaliana. Plant J. *83*, 991–1004.

Silva, P.A.G.C., Pereira, C.F., Dalebout, T.J., Spaan, W.J.M., and Bredenbeek, P.J. (2010). An RNA Pseudoknot Is Required for Production of Yellow Fever Virus Subgenomic RNA by the Host Nuclease XRN1. J. Virol. *84*, 11395.

Silvestri, L.S., Parilla, J.M., Morasco, B.J., Ogram, S.A., and Flanegan, J.B. (2006). Relationship between poliovirus negative-strand RNA synthesis and the length of the 3′ poly(A) tail. Virology *345*, 509–519.

Simon, A.E. (2015). 3'UTRs of carmoviruses. Virus Res. *206*, 27–36.

Simon, A.E., and Howell, S.H. (1986). The virulent satellite RNA of turnip crinkle virus has a major domain homologous to the 3' end of the helper virus genome. EMBO J. *5*, 3423–3428.

Simon, A.E., Roossinck, M.J., and Havelda, Z. (2004). Plant virus satellite and defective interfering RNAs: new paradigms for a new century. Annu. Rev. Phytopathol. *42*, 415–437.

Singh, R.N., and Dreher, T.W. (1998). Specific site selection in RNA resulting from a combination of nonspecific secondary structure and -CCR- boxes: initiation of minus strand synthesis by turnip yellow mosaic virus RNA-dependent RNA polymerase. RNA *4*, 1083–1095.

Slevin, M.K., Meaux, S., Welch, J.D., Bigler, R., Miliani de Marval, P.L., Su, W., Rhoads, R.E., Prins, J.F., and Marzluff, W.F. (2014). Deep Sequencing Shows Multiple Oligouridylations Are Required for 3′ to 5′ Degradation of Histone mRNAs on Polyribosomes. Mol. Cell *53*, 1020–1030.

Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J.M., and Watkins, N.J. (2013). Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. J. Cell Biol. *200*, 577–588.

Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2006). Polyadenylation of ribosomal RNA in human cells. Nucleic Acids Res. *34*, 2966–2975.

Smallwood, S., and Moyer, S.A. (1993). Promoter analysis of the vesicular stomatitis virus RNA polymerase. Virology *192*, 254–263.

Soma, N., Higashimoto, K., Imamura, M., Saitoh, A., Soejima, H., and Nagasaki, K. (2017). Long term survival of a patient with Perlman syndrome due to novel compound heterozygous missense mutations in RNB domain of DIS3L2. Am. J. Med. Genet. A *173*, 1077–1081.

Sõmera, M., Sarmiento, C., and Truve, E. (2015). Overview on Sobemoviruses and a Proposal for the Creation of the Family Sobemoviridae. Viruses *7*, 3076–3115.

Sõmera, M., Fargette, D., Hébrard, E., and Sarmiento, C. (2021). ICTV Virus Taxonomy Profile: Solemoviridae 2021. J. Gen. Virol. *102*.

Song, C., and Simon, A.E. (1995). Requirement of a 3'-terminal stem-loop in in vitro transcription by an RNA-dependent RNA polymerase. J. Mol. Biol. *254*, 6–14.

Song, M.-G., and Kiledjian, M. (2007). 3' Terminal oligo U-tract-mediated stimulation of decapping. RNA *13*, 2356–2365.

Song, J., Wang, X., Song, B., Gao, L., Mo, X., Yue, L., Yang, H., Lu, J., Ren, G., Mo, B., et al. (2019). Prevalent cytidylation and uridylation of precursor miRNAs in Arabidopsis. Nat. Plants *5*, 1260–1272.

Sorenson, R.S., Deshotel, M.J., Johnson, K., Adler, F.R., and Sieburth, L.E. (2018). Arabidopsis mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. Proc. Natl. Acad. Sci. U. S. A. *115*, E1485–E1494.

Souret, F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 Degrades mRNA in Arabidopsis and Its Substrates Include Selected miRNA Targets. Mol. Cell *15*, 173–183.

Stagno, J., Aphasizheva, I., Rosengarth, A., Luecke, H., and Aphasizhev, R. (2007a). UTP-bound and Apo structures of a minimal RNA uridylyltransferase. J. Mol. Biol. *366*, 882–899.

Stagno, J., Aphasizheva, I., Aphasizhev, R., and Luecke, H. (2007b). Dual role of the RNA substrate in selectivity and catalysis

by terminal uridylyl transferases. Proc. Natl. Acad. Sci. U. S. A. *104*, 14634–14639.

Stagno, J., Aphasizheva, I., Bruystens, J., Luecke Hartmut, H., and Aphasizhev, R. (2010). Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. J. Mol. Biol. *399*, 464–475.

Steckelberg, A.L., Vicens, Q., and Kieft, J.S. (2018a). Exoribonuclease-Resistant RNAs Exist within both Coding and Noncoding Subgenomic RNAs. MBio *9*.

Steckelberg, A.L., Akiyama, B.M., Costantino, D.A., Sit, T.L., Nix, J.C., and Kieft, J.S. (2018b). A folded viral noncoding RNA blocks host cell exoribonucleases through a conformationally dynamic RNA structure. Proc. Natl. Acad. Sci. U. S. A. *115*, 6404–6409.

Steil, B.P., and Barton, D.J. (2009). Cis-active RNA elements (CREs) and picornavirus RNA replication. Virus Res. *139*, 240–252.

Steil, B.P., Kempf, B.J., and Barton, D.J. (2010). Poly(A) at the 3' end of positive-strand RNA and VPg-linked poly(U) at the 5' end of negative-strand RNA are reciprocal templates during replication of poliovirus RNA. J. Virol. *84*, 2843–2858.

Stevens, A. (1978). An exoribonuclease from saccharomyces cerevisiae: Effect of modifications of 5′ end groups on the hydrolysis of substrates to 5′ mononucleotides. Biochem. Biophys. Res. Commun. *81*, 656–661.

Stupina, V., and Simon, A.E. (1997). Analysis in vivo of turnip crinkle virus satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. Virology *238*, 470–477.

Stupina, V.A., Meskauskas, A., Mccormack, J.C., Yingling, Y.G., Shapiro, B.A., Dinman, J.D., and Simon, A.E. (2008). The 3' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. RNA *14*, 2379–2393.

Stupina, V.A., Yuan, X., Meskauskas, A., Dinman, J.D., and Simon, A.E. (2011). Ribosome binding to a 5' translational enhancer is altered in the presence of the 3' untranslated region in cap-independent translation of turnip crinkle virus. J. Virol. *85*, 4638–4653.

Su, J.-L., Chen, P.-S., Johansson, G., and Kuo, M.-L. (2012). Function and regulation of let-7 family microRNAs. MicroRNA (Shariqah, United Arab Emirates) *1*, 34–39.

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature *508*, 66–71.

Sudo, H., Nozaki, A., Uno, H., Ishida, Y. ichi, and Nagahama, M. (2016). Interaction properties of human TRAMP-like proteins and their role in pre-rRNA 5'ETS turnover. FEBS Lett. *590*, 2963–2972.

Suematsu, T., Zhang, L., Aphasizheva, I., Monti, S., Huang, L., Wang, Q., Costello, C.E., and Aphasizhev, R. (2016). Antisense Transcripts Delimit Exonucleolytic Activity of the Mitochondrial 3' Processome to Generate Guide RNAs. Mol. Cell *61*, 364–378.

Sulzinski, M.A., and Zaitlin, M. (1982). Tobacco mosaic virus replication in resistant and susceptible plants: in some resistant species virus is confined to a small number of initially infected cells. Virology *121*, 12–19.

Sun, X., and Simon, A.E. (2006). A cis-replication element functions in both orientations to enhance replication of Turnip crinkle virus. Virology *352*, 39–51.

Sun, M., Schwalb, B., Pirkl, N., Maier, K.C., Schenk, A., Failmezger, H., Tresch, A., and Cramer, P. (2013). Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. Mol. Cell *52*, 52–62.

Swenson, K.G. (1952). Aphid transmission of a strain of Alfalfa mosaic virus. Phytopathology *42*, 261–262.

Tacahashi, Y., and Uyeda, I. (1999). Restoration of the 3' end of potyvirus RNA derived from Poly(A)-deficient infectious cDNA clones. Virology *265*, 147–152.

Tacke, E., Prufer, D., Salamini, F., and Rohde, W. (1990). Characterization of a potato leafroll luteovirus subgenomic RNA: Differential expression by internal translation initiation and UAG suppression. J. Gen. Virol. *71*, 2265–2272.

Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. Mol. Cell *51*, 539–551.

Tan, D., Marzluff, W.F., Dominski, Z., and Tong, L. (2013). Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3'hExo ternary complex. Science *339*, 318–321.

Tang, D., Wang, G., and Zhou, J.M. (2017). Receptor Kinases in Plant-Pathogen Interactions: More Than Pattern Recognition. Plant Cell *29*, 618–637.

Tarun, J., Wells, S.E., Deardorff, J.A., and Sachs, A.B. (1997). Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. Proc. Natl. Acad. Sci. U. S. A. *94*, 9046.

Tavert-Roudet, G., Anne, A., Barra, A., Chovin, A., Demaille, C., and Michon, T. (2017). The Potyvirus Particle Recruits the Plant Translation Initiation Factor eIF4E by Means of the VPg covalently Linked to the Viral RNA. Mol. Plant. Microbe. Interact. *30*, 754–762.

Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. Nature *404*, 515–518.

Thivierge, K., Cotton, S., Dufresne, P.J., Mathieu, I., Beauchemin, C., Ide, C., Fortin, M.G., and Laliberté, J.F. (2008). Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. Virology *377*, 216–225.

Thomas, C.L., Leh, V., Lederer, C., and Maule, A.J. (2003). Turnip crinkle virus coat protein mediates suppression of RNA

silencing in Nicotiana benthamiana. Virology *306*, 33–41.

Thomas, M.P., Liu, X., Whangbo, J., McCrossan, G., Sanborn, K.B., Basar, E., Walch, M., and Lieberman, J. (2015). Apoptosis Triggers Specific, Rapid, and Global mRNA Decay with 3' Uridylated Intermediates Degraded by DIS3L2. Cell Rep. *11*, 1079–1089.

Thomas, P.E., Wu, X., Liu, M., Gaffney, B., Ji, G., Li, Q.Q., and Hunt, A.G. (2012). Genome-wide control of polyadenylation site choice by CPSF30 in Arabidopsis. Plant Cell *24*, 4376–4388.

Thompson, J.R., Buratti, E., de Wispelaere, M., and Tepfer, M. (2008). Structural and functional characterization of the 5′ region of subgenomic RNA5 of cucumber mosaic virus. J. Gen. Virol. *89*, 1729–1738.

Thompson, J.R., Dasgupta, I., Fuchs, M., Iwanami, T., Karasev, A. V., Petrzik, K., Sanfaçon, H., Tzanetakis, I., van Der Vlugt, R., Wetzel, T., et al. (2017). ICTV Virus Taxonomy Profile: Secoviridae. J. Gen. Virol. *98*, 529–531.

Thornton, J.E., Chang, H.M., Piskounova, E., and Gregory, R.I. (2012). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA *18*, 1875–1885.

Thornton, J.E., Du, P., Jing, L., Sjekloca, L., Lin, S., Grossi, E., Sliz, P., Zon, L.I., and Gregory, R.I. (2014). Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). Nucleic Acids Res. *42*, 11777–11791.

Till, D.D., Linz, B., Seago, J.E., Elgar, S.J., Marujo, P.E., Elias, M.D.L., Arraiano, C.M., McClellan, J.A., McCarthy, J.E.G., and Newbury, S.F. (1998). Identification and developmental expression of a 5'-3' exoribonuclease from Drosophila melanogaster. Mech. Dev. *79*, 51–55.

Tomar, S., Hardy, R.W., Smith, J.L., and Kuhn, R.J. (2006). Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. J. Virol. *80*, 9962–9969.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., et al. (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. EMBO J. *29*, 2342–2357.

Torabi, S.F., Chen, Y.L., Zhang, K., Wang, J., DeGregorio, S.J., Vaidya, A.T., Su, Z., Pabit, S.A., Chiu, W., Pollack, L., et al. (2021a). Structural analyses of an RNA stability element interacting with poly(A). Proc. Natl. Acad. Sci. U. S. A. *118*.

Torabi, S.F., Vaidya, A.T., Tycowski, K.T., DeGregorio, S.J., Wang, J., Shu, M. Di, Steitz, T.A., and Steitz, J.A. (2021b). RNA stabilization by a poly(A) tail 3'-end binding pocket and other modes of poly(A)-RNA interaction. Science *371*.

Torrance, L., Cowan, G.H., Gillespie, T., Ziegler, A., and Lacomme, C. (2006). Barley stripe mosaic virus-encoded proteins triple-gene block 2 and γb localize to chloroplasts in virus-infected monocot and dicot plant, revealing hitherto-unknown roles in virus replication. J. Gen. Virol. *87*, 2403–2411.

Towler, B.P., Jones, C.I., Harper, K.L., Waldron, J.A., and Newbury, S.F. (2016). A novel role for the 3'-5' exoribonuclease Dis3L2 in controlling cell proliferation and tissue growth. RNA Biol. *13*, 1286–1299.

Trippe, R., Sandrock, B., and Benecke, B.J. (1998). A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. Nucleic Acids Res. *26*, 3119–3126.

Trippe, R., Richly, H., and Benecke, B.-J. (2003). Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase. Eur. J. Biochem. *270*, 971–980.

Trippe, R., Guschina, E., Hossbach, M., Urlaub, H., Lührmann, R., and Benecke, B.-J. (2006). Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA *12*, 1494–1504.

Tritschler, F., Eulalio, A., Helms, S., Schmidt, S., Coles, M., Weichenrieder, O., Izaurralde, E., and Truffault, V. (2008). Similar modes of interaction enable Trailer Hitch and EDC3 to associate with DCP1 and Me31B in distinct protein complexes. Mol. Cell. Biol. *28*, 6695–6708.

Trost, G., Vi, S.L., Czesnick, H., Lange, P., Holton, N., Giavalisco, P., Zipfel, C., Kappel, C., and Lenhard, M. (2014). Arabidopsis poly(A) polymerase PAPS1 limits founder-cell recruitment to organ primordia and suppresses the salicylic acid-independent immune response downstream of EDS1/PAD4. Plant J. *77*, 688–699.

Tsai, C.H., and Dreher, T.W. (1991). Turnip yellow mosaic virus RNAs with anticodon loop substitutions that result in decreased valylation fail to replicate efficiently. J. Virol. *65*, 3060–3067.

Tsai, C.-H., Cheng, C.-P., Peng, C.-W., Lin, B.-Y., Lin, N.-S., and Hsu, Y.-H. (1999). Sufficient length of a poly(A) tail for the formation of a potential pseudoknot is required for efficient replication of bamboo mosaic potexvirus RNA. J. Virol. *73*, 2703–2709.

Tsuzuki, M., Motomura, K., Kumakura, N., and Takeda, A. (2017). Interconnections between mRNA degradation and RDR-dependent siRNA production in mRNA turnover in plants. J. Plant Res. *130*, 211–226.

Tu, B., Liu, L., Xu, C., Zhai, J., Li, S., Lopez, M.A., Zhao, Y., Yu, Y., Ramachandran, V., Ren, G., et al. (2015). Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in Arabidopsis. PLoS Genet. *11*, e1005119.

Tuck, A.C., Rankova, A., Arpat, A.B., Liechti, L.A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., and Bühler, M. (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. Mol. Cell *77*, 1222-1236.e13.

Tudek, A., Lloret-Llinares, M., and Heick Jensen, T. (2018). The multitasking polyA tail: nuclear RNA maturation, degradation and export. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *373*.

Tudek, A., Krawczyk, P.S., Mroczek, S., Tomecki, R., Turtola, M., Matylla-Kulińska, K., Jensen, T.H., and Dziembowski, A. (2021). Global view on the metabolism of RNA poly(A) tails in yeast Saccharomyces cerevisiae. Nat. Commun. 2021 121 *12*, 1–14.

Turina, M., Desvoyes, B., and Scholthof, K.B.G. (2000). A Gene Cluster Encoded by Panicum Mosaic Virus Is Associated with

Virus Movement. Virology *266*, 120–128.

Turnbull-Ross, A.D., Reavy, B., Mayo, M.A., and Murant, A.F. (1992). The nucleotide sequence of parsnip yellow fleck virus: a plant picorna-like virus. J. Gen. Virol. *73 (Pt 12)*, 3203–3211.

Tvarogová, J., Madhugiri, R., Bylapudi, G., Ferguson, L.J., Karl, N., and Ziebuhr, J. (2019). Identification and characterization of a human coronavirus 229E nonstructural protein 8-associated RNA 3'-terminal adenylyltransferase activity. J. Virol. JVI.00291-19.

Tzanetakis, I.E., Tsai, C.H., Martin, R.R., and Dreher, T.W. (2009). A tymovirus with an atypical 3'-UTR illuminates the possibilities for 3'-UTR evolution. Virology *392*, 238–245.

Uchida, N., Hoshino, S.I., and Katada, T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. J. Biol. Chem. *279*, 1383–1391.

Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z., and Vanacova, S. (2013). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA *19*, 1632–1638.

Ustianenko, D., Pasulka, J., Feketova, Z., Bednarik, L., Zigackova, D., Fortova, A., Zavolan, M., and Vanacova, S. (2016). TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. EMBO J. *35*, 2179–2191.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. *20*, 515–524.

Valli, A.A., Gallo, A., Rodamilans, B., López-Moya, J.J., and García, J.A. (2018). The HCPro from the Potyviridae family: an enviable multitasking Helper Component that every virus would like to have. Mol. Plant Pathol. *19*, 744–763.

Vaňáčová, Š., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol. *3*, 0986–0997.

Várallyay, É., Válóczi, A., Ágyi, Á., Burgyán, J., and Havelda, Z. (2010). Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. EMBO J. *29*, 3507–3519.

Venkataraman, S., Prasad, B.V.L.S., and Selvarajan, R. (2018). RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. Viruses *10*.

Verchot, J. (2022). Potato virus X: A global potato-infecting virus and type member of the Potexvirus genus. Mol. Plant Pathol. *23*, 315–320.

Vi, S.L., Trost, G., Lange, P., Czesnick, H., Rao, N., Lieber, D., Laux, T., Gray, W.M., Manley, J.L., Groth, D., et al. (2013). Target specificity among canonical nuclear poly(A) polymerases in plants modulates organ growth and pathogen response. Proc. Natl. Acad. Sci. U. S. A. *110*, 13994–13999.

Vidya, E., and Duchaine, T.F. (2022). Eukaryotic mRNA Decapping Activation. Front. Genet. *13*.

Vigh, M.L., Bressendorff, S., Thieffry, A., Arribas-Hernández, L., and Brodersen, P. (2022). Nuclear and cytoplasmic RNA exosomes and PELOTA1 prevent miRNA-induced secondary siRNA production in Arabidopsis. Nucleic Acids Res. *50*, 1396–1415.

Vijayakumar, A., Park, A., and Steitz, J.A. (2022). Modulation of mRNA 3′-End Processing and Transcription Termination in Virus-Infected Cells. Front. Immunol. *13*, 402.

Vogel, F., Hofius, D., Paulus, K.E., Jungkunz, I., and Sonnewald, U. (2011). The second face of a known player: Arabidopsis silencing suppressor AtXRN4 acts organ-specifically. New Phytol. *189*, 484–493.

Volchkova, V.A., Vorac, J., Repiquet-Paire, L., Lawrence, P., and Volchkov, V.E. (2015). Ebola Virus GP Gene Polyadenylation Versus RNA Editing. J. Infect. Dis. *212 Suppl 2*, S191–S198.

Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., and Dehesh, K. (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. Plant Physiol. *152*, 866–875.

Walsh, J.A., and Jenner, C.E. (2002). Turnip mosaic virus and the quest for durable resistance. Mol. Plant Pathol. *3*, 289–300.

Walter, P., and Blobel, G. (1982). Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature *299*, 691–698.

Walters, R.W., Matheny, T., Mizoue, L.S., Rao, B.S., Muhlrad, D., and Parker, R. (2017). Identification of NAD+ capped mRNAs in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *114*, 480–485.

Wang, C.-Y., Chen, W.-L., and Wang, S.-W. (2013a). Pdc1 functions in the assembly of P bodies in Schizosaccharomyces pombe. Mol. Cell. Biol. *33*, 1244–1253.

Wang, L., Nam, Y., Lee, A.K., Yu, C., Roth, K., Chen, C., Ransey, E.M., and Sliz, P. (2017). LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. Cell Rep. *18*, 2664–2675.

Wang, X., Zhang, S., Dou, Y., Zhang, C., Chen, X., Yu, B., and Ren, G. (2015). Synergistic and Independent Actions of Multiple Terminal Nucleotidyl Transferases in the 3' Tailing of Small RNAs in Arabidopsis. PLOS Genet. *11*, e1005091.

Wang, X., Kong, W., Wang, Y., Wang, J., Zhong, L., Lao, K., Dong, X., Zhang, D., Huang, H., Mo, B., et al. (2022). Uridylation and the SKI complex orchestrate the Calvin cycle of photosynthesis through RNA surveillance of *TKL1* in Arabidopsis. Proc. Natl. Acad. Sci. *119*.

Wang, Y., Li, S., Zhao, Y., You, C., Le, B., Gong, Z., Mo, B., Xia, Y., and Chen, X. (2019). NAD +-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated. Proc. Natl. Acad. Sci. U. S. A. *116*, 12094–12102.

Wang, Y., Weng, C., Chen, X., Zhou, X., Huang, X., Yan, Y., and Zhu, C. (2020). CDE-1 suppresses the production of risiRNA by coupling polyuridylation and degradation of rRNA. BMC Biol. *18*.

Wang, Z., Jiao, X., Carr-Schmid, A., and Kiledjian, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. Proc. Natl. Acad. Sci. U. S. A. *99*, 12663–12668.

Wang, Z., Qiu, Y., Liu, Y., Qi, N., Si, J., Xia, X., Wu, D., Hu, Y., and Zhou, X. (2013b). Characterization of a nodavirus replicase revealed a de novo initiation mechanism of RNA synthesis and terminal nucleotidyltransferase activity. J. Biol. Chem. *288*, 30785–30801.

Warkocki, Z., Liudkovska, V., Gewartowska, O., Mroczek, S., and Dziembowski, A. (2018a). Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *373*.

Warkocki, Z., Krawczyk, P.S., Adamska, D., Bijata, K., Garcia-Perez, J.L., and Dziembowski, A. (2018b). Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. Cell *174*, 1537-1548.e29.

Wasmuth, E. V., and Lima, C.D. (2012). Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. Mol. Cell *48*, 133–144.

Wasmuth, E. V., and Lima, C.D. (2016). The Rrp6 C-terminal domain binds RNA and activates the nuclear RNA exosome. Nucleic Acids Res. *45*, 846–860.

Weaver, B.P., Zabinsky, R., Weaver, Y.M., Lee, E.S., Xue, D., and Han, M. (2014). CED-3 caspase acts with miRNAs to regulate non-apoptotic gene expression dynamics for robust development in C. elegans. Elife *3*, 1–22.

Webster, M.W., Chen, Y.-H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Mol. Cell *70*, 1089-1100.e8.

Webster, M.W., Stowell, J.A.W., and Passmore, L.A. (2019). RNA-binding proteins distinguish between similar sequence motifs to promote targeted deadenylation by Ccr4-Not. Elife *8*.

Wegert, J., Ishaque, N., Vardapour, R., Geörg, C., Gu, Z., Bieg, M., Ziegler, B., Bausenwein, S., Nourkami, N., Ludwig, N., et al. (2015). Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors. Cancer Cell *27*, 298–311.

Wei, T., and Wang, A. (2008). Biogenesis of Cytoplasmic Membranous Vesicles for Plant Potyvirus Replication Occurs at Endoplasmic Reticulum Exit Sites in a COPI- and COPII-Dependent Manner. J. Virol. *82*, 12252.

Wei, C.M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. Cell *4*, 379–386.

Weick, E.M., and Lima, C.D. (2021). RNA helicases are hubs that orchestrate exosome-dependent 3'-5' decay. Curr. Opin. Struct. Biol. *67*, 86–94.

Welch, J.D., Slevin, M.K., Tatomer, D.C., Duronio, R.J., Prins, J.F., and Marzluff, W.F. (2015). EnD-Seq and AppEnD: sequencing 3' ends to identify nontemplated tails and degradation intermediates. RNA *21*, 1375–1389.

Werner, M., Purta, E., Kaminska, K.H., Cymerman, I.A., Campbell, D.A., Mittra, B., Zamudio, J.R., Sturm, N.R., Jaworski, J., and Bujnicki, J.M. (2011). 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. Nucleic Acids Res. *39*, 4756–4768.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522–525.

Wilusz, J.E., Whipple, J.M., Phizicky, E.M., and Sharp, P.A. (2011). tRNAs marked with CCACCA are targeted for degradation. Science *334*, 817–821.

Wittmann, S., Chatel, H., Fortin, M.G., and Laliberté, J.F. (1997). Interaction of the viral protein genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso) 4E of Arabidopsis thaliana using the yeast two-hybrid system. Virology *234*, 84–92.

Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C. V, Bycroft, M., Stewart, M., et al. (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. EMBO J. *33*, 1514–1526.

van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135–148.

Workman, R.E., Tang, A.D., Tang, P.S., Jain, M., Tyson, J.R., Razaghi, R., Zuzarte, P.C., Gilpatrick, T., Payne, A., Quick, J., et al. (2019). Nanopore native RNA sequencing of a human poly(A) transcriptome. Nat. Methods *16*, 1297.

Wu, B., and White, K.A. (2007). Uncoupling RNA virus replication from transcription via the polymerase: functional and evolutionary insights. EMBO J. *26*, 5120–5130.

Wu, B., Grigull, J., Ore, M.O., Morin, S., and White, K.A. (2013a). Global organization of a positive-strand RNA virus genome. PLoS Pathog. *9*.

Wu, H.Y., Ke, T.Y., Liao, W.Y., and Chang, N.Y. (2013b). Regulation of coronaviral poly(A) tail length during infection. PLoS One *8*.

Wu, M., Reuter, M., Lilie, H., Liu, Y., Wahle, E., and Song, H. (2005). Structural insight into poly(A) binding and catalytic mechanism of human PARN. EMBO J. *24*, 4082.

Wu, S., Li, X., and Wang, G. (2021). tRNA-like structures and their functions. FEBS J.

Wu, W., Wang, Z., Xia, H., Liu, Y., Qiu, Y., Liu, Y., Hu, Y., and Zhou, X. (2014). Flock House Virus RNA Polymerase Initiates RNA Synthesis De Novo and Possesses a Terminal Nucleotidyl Transferase Activity. PLoS One *9*, e86876.

Wu, X., Valli, A., García, J.A., Zhou, X., and Cheng, X. (2019). The Tug-of-War between Plants and Viruses: Great Progress and Many Remaining Questions. Viruses 2019, Vol. 11, Page 203 *11*, 203.

Wyman, S.K., Knouf, E.C., Parkin, R.K., Fritz, B.R., Lin, D.W., Dennis, L.M., Krouse, M.A., Webster, P.J., and Tewari, M. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome Res. *21*, 1450–1461.

Xu, J., and Chua, N.H. (2009). Arabidopsis decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. Plant Cell *21*, 3270–3279.

Xu, K., and Nagy, P.D. (2014). Expanding use of multi-origin subcellular membranes by positive-strand RNA viruses during replication. Curr. Opin. Virol. *9*, 119–126.

Xu, J., Yang, J.Y., Niu, Q.W., and Chua, N.H. (2006). Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. Plant Cell *18*, 3386–3398.

Xu, K., Lin, J., Zandi, R., Roth, J.A., and Ji, L. (2016). MicroRNA-mediated target mRNA cleavage and 3′-uridylation in human cells. Sci. Rep. *6*, 30242.

Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.Y.A., and Shyu, A. Bin (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat. Struct. Mol. Biol. *12*, 1054–1063.

Yamashita, S., Takagi, Y., Nagaike, T., and Tomita, K. (2017). Crystal structures of U6 snRNA-specific terminal uridylyltransferase. Nat. Commun. *8*.

Yang, A., Shao, T.J., Bofill-De Ros, X., Lian, C., Villanueva, P., Dai, L., and Gu, S. (2020). AGO-bound mature miRNAs are oligouridylated by TUTs and subsequently degraded by DIS3L2. Nat. Commun. *11*.

Yang, H., Gou, X., He, K., Xi, D., Du, J., Lin, H., and Li, J. (2010). BAK1 and BKK1 in Arabidopsis thaliana confer reduced susceptibility to turnip crinkle virus. Eur. J. Plant Pathol. *127*, 149–156.

Yang, W., Lee, J.Y., and Nowotny, M. (2006a). Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. Mol. Cell *22*, 5–13.

Yang, Z., Ebright, Y.W., Yu, B., and Chen, X. (2006b). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. Nucleic Acids Res. *34*, 667–675.

Yates, L.A., Fleurdépine, S., Rissland, O.S., De Colibus, L., Harlos, K., Norbury, C.J., and Gilbert, R.J.C. (2012). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nat. Struct. Mol. Biol. *19*, 782–787.

Yates, L.A., Durrant, B.P., Fleurdépine, S., Harlos, K., Norbury, C.J., and Gilbert, R.J.C. (2015). Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. Nucleic Acids Res. *43*, 2968–2979.

Yeam, I., Cavatorta, J.R., Ripoll, D.R., Kang, B.C., and Jahn, M.M. (2007). Functional dissection of naturally occurring amino acid substitutions in eIF4E that confers recessive potyvirus resistance in plants. Plant Cell *19*, 2913–2928.

Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N. (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Mol. Cell *70*, 1081-1088.e5.

Yot, P., Pinck, M., Haenni, A.L., Duranton, H.M., and Chapeville, F. (1970). Valine-specific tRNA-like structure in turnip yellow mosaic virus RNA. Proc. Natl. Acad. Sci. U. S. A. *67*, 1345–1352.

Yu, S., and Kim, V.N. (2020). A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. Nat. Rev. Mol. Cell Biol. 2020 219 *21*, 542–556.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. Science *307*, 932–935.

Yu, X., Willmann, M.R., Anderson, S.J., and Gregory, B.D. (2016). Genome-Wide Mapping of Uncapped and Cleaved Transcripts Reveals a Role for the Nuclear mRNA Cap-Binding Complex in Cotranslational RNA Decay in Arabidopsis. Plant Cell *28*, 2385–2397.

Yu, Y., Ji, L., Le, B.H., Zhai, J., Chen, J., Luscher, E., Gao, L., Liu, C., Cao, X., Mo, B., et al. (2017). ARGONAUTE10 promotes the degradation of miR165/6 through the SDN1 and SDN2 exonucleases in Arabidopsis. PLoS Biol. *15*.

Yuan, X., Shi, K., Meskauskas, A., and Simon, A.E. (2009). The 3' end of Turnip crinkle virus contains a highly interactive structure including a translational enhancer that is disrupted by binding to the RNA-dependent RNA polymerase. RNA *15*, 1849–1864.

Yuan, X., Shi, K., Young, M.Y.L., and Simon, A.E. (2010). The terminal loop of a 3' proximal hairpin plays a critical role in replication and the structure of the 3' region of Turnip crinkle virus. Virology *402*, 271–280.

Yuan, X., Shi, K., and Simon, A.E. (2012). A local, interactive network of 3' RNA elements supports translation and replication of Turnip crinkle virus. J. Virol. *86*, 4065–4081.

Zeng, W., Dai, X., Sun, J., Hou, Y., Ma, X., Cao, X., Zhao, Y., and Cheng, Y. (2019). Modulation of Auxin Signaling and Development by Polyadenylation Machinery. Plant Physiol. *179*, 686–699.

Zhai, J., Zhao, Y., Simon, S.A., Huang, S., Petsch, K., Arikit, S., Pillay, M., Ji, L., Xie, M., Cao, X., et al. (2013). Plant microRNAs display differential 3' truncation and tailing modifications that are ARGONAUTE1 dependent and conserved across species. Plant Cell *25*, 2417–2428.

Zhang, C.W., Liu, Q., Zeng, Q., Huang, W.T., Wang, Q., and Cheng, Y.Q. (2020a). p24 G1 Encoded by Grapevine Leafroll-

Associated Virus 1 Suppresses RNA Silencing and Elicits Hypersensitive Response-Like Necrosis in Nicotiana Species. Viruses *12*.

Zhang, E., Khanna, V., Dacheux, E., Namane, A., Doyen, A., Gomard, M., Turcotte, B., Jacquier, A., and Fromont-Racine, M. (2019a). A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. EMBO J. *38*, e100640.

Zhang, G., Zhang, J., and Simon, A.E. (2004). Repression and Derepression of Minus-Strand Synthesis in a Plus-Strand RNA Virus Replicon. J. Virol. *78*, 7619–7633.

Zhang, G., Zhang, J., George, A.T., Baumstark, T., and Simon, A.E. (2006a). Conformational changes involved in initiation of minus-strand synthesis of a virus-associated RNA. RNA *12*, 147–162.

Zhang, H., Tang, K., Qian, W., Duan, C.G., Wang, B., Zhang, H., Wang, P., Zhu, X., Lang, Z., Yang, Y., et al. (2014). An Rrp6-like Protein Positively Regulates Non-coding RNA Levels and DNA Methylation in Arabidopsis. Mol. Cell *54*, 418.

Zhang, H., Zhong, H., Zhang, S., Shao, X., Ni, M., Cai, Z., Chen, X., and Xia, Y. (2019b). NAD tagSeq reveals that NAD +-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. *116*, 12072–12077.

Zhang, J., Zhang, G., McCormack, J.C., and Simon, A.E. (2006b). Evolution of virus-derived sequences for high-level replication of a subviral RNA. Virology *351*, 476–488.

Zhang, J., Zhang, G., Guo, R., Shapiro, B.A., and Simon, A.E. (2006c). A Pseudoknot in a Preactive Form of a Viral RNA Is Part of a Structural Switch Activating Minus-Strand Synthesis. J. Virol. *80*, 9181–9191.

Zhang, S., Li, R., Zhang, L., Chen, S., Xie, M., Yang, L., Xia, Y., Foyer, C.H., Zhao, Z., and Lam, H.M. (2020b). New insights into Arabidopsis transcriptome complexity revealed by direct sequencing of native RNAs. Nucleic Acids Res. *48*, 7700–7711.

Zhang, S., Sun, R., Perdoncini Carvalho, C., Han, J., Zheng, L., and Qu, F. (2021). Replication-Dependent Biogenesis of Turnip Crinkle Virus Long Noncoding RNAs. J. Virol. *95*.

Zhang, W., Murphy, C., and Sieburth, L.E. (2010). Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. Proc. Natl. Acad. Sci. U. S. A. *107*, 15981–15985.

Zhang, X., Zhang, X., Singh, J., Li, D., and Qu, F. (2012). Temperature-dependent survival of Turnip crinkle virus-infected arabidopsis plants relies on an RNA silencing-based defense that requires dcl2, AGO2, and HEN1. J. Virol. *86*, 6847–6854.

Zhang, X., Devany, E., Murphy, M.R., Glazman, G., Persaud, M., and Kleiman, F.E. (2015a). PARN deadenylase is involved in miRNA-dependent degradation of TP53 mRNA in mammalian cells. Nucleic Acids Res. *43*, 10925–10938.

Zhang, X., Zhu, Y., Liu, X., Hong, X., Xu, Y., Zhu, P., Shen, Y., Wu, H., Ji, Y., Wen, X., et al. (2015b). Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. Science *348*, 120–123.

Zhang, Z., Hu, F., Sung, M.W., Shu, C., Castillo-González, C., Koiwa, H., Tang, G., Dickman, M., Li, P., and Zhang, X. (2017). RISC-interacting clearing 3'- 5' exoribonucleases (RICEs) degrade uridylated cleavage fragments to maintain functional RISC in Arabidopsis thaliana. Elife *6*, e24466.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012). The Arabidopsis Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation. Curr. Biol. *22*, 689–694.

Zheng, H., Lee, H.A., Palese, P., and García-Sastre, A. (1999). Influenza A virus RNA polymerase has the ability to stutter at the polyadenylation site of a viral RNA template during RNA replication. J. Virol. *73*, 5240–5243.

Zhou, X., Feng, X., Mao, H., Li, M., Xu, F., Hu, K., and Guang, S. (2017). RdRP-synthesized antisense ribosomal siRNAs silence pre-rRNA via the nuclear RNAi pathway. Nat. Struct. Mol. Biol. *24*, 258–269.

Zhu, H.Y., Ling, K.S., Goszczynski, D.E., McFerson, J.R., and Gonsalves, D. (1998). Nucleotide sequence and genome organization of grapevine leafroll-associated virus-2 are similar to beet yellows virus, the closterovirus type member. J. Gen. Virol. *79 (Pt 5)*, 1289–1298.

Zhu, L., Hu, Q., Cheng, L., Jiang, Y., Lv, M., Liu, Y., Li, F., Shi, Y., and Gong, Q. (2020). Crystal structure of Arabidopsis terminal uridylyl transferase URT1. Biochem. Biophys. Res. Commun. *524*, 490–496.

Zigáčková, D., and Vaňáčová, Š. (2018). The role of 3′ end uridylation in RNA metabolism and cellular physiology. Philos. Trans. R. Soc. B Biol. Sci. *373*, 20180171.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev. *31*, 88–100.

Zuber, H., Scheer, H., Ferrier, E., Sement, F.M., Mercier, P., Stupfler, B., and Gagliardi, D. (2016). Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis. Cell Rep. *14*, 2707–2717.

Zuber, H., Scheer, H., Joly, A.-C., and Gagliardi, D. (2018). Respective Contributions of URT1 and HESO1 to the Uridylation of 5′ Fragments Produced From RISC-Cleaved mRNAs. Front. Plant Sci. *9*, 1438.

Zuo, X., Wang, J., Yu, P., Eyler, D., Xu, H., Starich, M.R., Tiede, D.M., Simon, A.E., Kasprzak, W., Schwieters, C.D., et al. (2010). Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3' UTR of turnip crinkle virus. Proc. Natl. Acad. Sci. U. S. A. *107*, 1385–1390.

Zuo, Z., Roux, M., Rodriguez, E., and Petersen, M. (2022). mRNA Decapping Factors LSM1 and PAT Paralogs Are Involved in Turnip Mosaic Virus Viral Infection. Mol. Plant-Microbe Interact. *35*, 125–130.

Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., et al. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA

dependent on the RNA sensor Mda5. Nat. Immunol. *12*, 137–143.

Anne Caroline JOLY

Diversity of the mechanisms related to viral RNA uridylation in plants

Résumé

Les phytovirus sont responsables d'importants dommages sur les cultures. La compréhension des mécanismes liés à l'infection virale est donc cruciale. Ces dernières années, l'uridylation des ARN est apparue comme un acteur clé de la régulation post-transcriptionnelle des ARN cellulaire mais également proposée comme un nouveau mécanisme antiviral chez les métazoaires. Pourtant, les connaissances sur les processus moléculaires impliqués dans l''uridylation des ARN viraux restent rares, en particulier pour les virus de plantes. Cette étude fournit des informations sur l'uridylation des ARN viraux chez les plantes pour des espèces représentatives des principales familles virales. Nos données révèlent que les ARN de phytovirus présentent des profils d'uridylation variés, ce qui suggère que des mécanismes moléculaires distincts pourraient être à l'œuvre. Mes résultats montrent l'implication des uridylyltransférases terminales (TUTases) de l'hôte dans l'uridylation des ARN du *Turnip mosaic virus* and du *Turnip crinkle virus*, tandis que d'autres facteurs, potentiellement des facteurs viraux, sont responsables du haut niveau de mono-uridylation observé pour les ARN du *Grapevine fanleaf virus*. Je montre également que l'accumulation du *Turnip mosaic virus* est réduite lors de la surexpression de HESO1, une uridylyl transférase terminale (TUTase) chez Arabidopsis. Ces travaux ouvrent la voie à une meilleure compréhension des mécanismes moléculaires sous-jacents de l'uridylation des ARN viraux chez les plantes.

Mot clés : phytovirus, ARN, uridylation, TUTases, 3'RACE-seq, *Arabidopsis thaliana*

Abstract

Phytoviruses cause significant crop damages worldwide. Therefore, the understanding of the mechanisms related to viral infection is crucial. In the last few years, RNA uridylation emerged as a new potential antiviral defense mechanism in metazoans. Yet, the knowledge of the molecular processes involved in viral RNA uridylation remains scarce, especially for plant viruses. This study provides a global insight on viral RNA uridylation in plants for species representative of the main families of single-stranded RNA phytoviruses. Our data reveal an extreme diversity of uridylation patterns across phytoviral RNAs, suggesting that distinct molecular mechanisms could be in place. In line with this, my results show the involvement of the host terminal uridylyltransferases (TUTases) in the RNA uridylation for the *Turnip mosaic virus* and *Turnip crinkle virus*, while other factors, possibly viral factors, are responsible for the high mono-uridylation level observed for the *Grapevine fanleaf virus* RNAs. Interestingly, I also show that the accumulation of the TuMV is reduced upon the overexpression of HESO1, an Arabidopsis TUTase. This work paves the way for the better understanding of the molecular mechanisms underlying viral RNA uridylation in plants.

Keywords: positive-strand RNA phytovirus, RNA, uridylation, TUTases, 3'RACE-seq, *Arabidopsis thaliana*