Proteome expansion in the Potyviridae evolutionary radiation

Fabio Pasin 1,2,*, José-Antonio Darós 1, Ioannis E. Tzanetakis 3

1Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de Investigaciones Científicas-Universitat Politècnica de València (CSIC-UPV), 46011 Valencia, Spain
2School of Science, University of Padova, 35121 Padova, Italy
3Department of Entomology and Plant Pathology, Division of Agriculture, University of Arkansas System, 72701 Fayetteville, AR, USA

*Corresponding author: Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de Investigaciones Científicas-Universitat Politècnica de València (CSIC-UPV), UPV Building 8E, Ingeniero Fausto Elio, 46011 Valencia, Spain. E-mail: f.pasin@csic.es

One sentence summary: Potyviridae represents one of the most notable evolutionary radiations among RNA viruses; its species richness is accompanied by large genomic layout diversity and expansion of the encoded protein modules.

Editor: Guo Hui-Shan

Abstract

Potyviridae, the largest family of known RNA viruses (realm Riboviria), belongs to the picorna-like supergroup and has important agricultural and ecological impacts. Potyvirid genomes are translated into polyproteins, which are in turn hydrolyzed to release mature products. Recent sequencing efforts revealed an unprecedented number of potyvirids with a rich variability in gene content and genomic layouts. Here, we review the heterogeneity of non-core modules that expand the structural and functional diversity of the potyvirid proteomes. We provide a family-wide classification of P1 proteinases into the functional Types A and B, and discuss pretty interesting sweet potato potyviral ORF (PISPO), putative zinc fingers, and alkylation B (AlkB)—non-core modules found within P1 cistrons. The atypical inosine triphosphate pyrophosphatase (ITPase/HAM1), as well as the pseudo tobacco mosaic virus-like coat protein (TMV-like CP) are discussed alongside homologs of unrelated virus taxa. Family-wide abundance of the multitasking helper component proteinase (HC-pro) is revised. Functional connections between non-core modules are highlighted to support host niche adaptation and immune evasion as main drivers of the Potyviridae evolutionary radiation. Potential biotechnological and synthetic biology applications of potyvirid leader proteinases and non-core modules are finally explored.

Keywords: Potyviridae, virus comparative genomics, non-core proteome module, evolutionary radiation, host adaptation, immune evasion

Introduction

Understanding the origin and evolution of viruses is complex, yet it is fundamental to fully realize the ecological, agricultural and medical impact of the virosphere (Jones and Naidu 2019, Zimmerman et al. 2020, Holmes et al. 2021, Liang and Bushman 2021). Plant virus diseases are major threats to food security; they occur worldwide and greatly affect developing countries (Jones and Naidu 2019, Savary et al. 2019). Conceptual frameworks rationalize the polyphyletic origins and evolution of the virome, as well as its ecological impact on crops and wild species (Lefevre et al. 2019, Dolja, Kruovic and Koonin 2020). Genomic resources for plant viruses have increased in the past four decades (Pasin, Menzel and Darós 2019), but our knowledge of plant virus evolution and host adaptation mechanisms is nonetheless incomplete.

The plant-infecting Potyviridae is the largest RNA virus family (realm Riboviria) (Fig. 1A). The most recent virus taxonomy based on phylogenomic analyses places the family within the phylum Psiviricota (Fig. 1B), which comprises of members of the former picorna-like supergroup (Koonin et al. 2020). Potyvirid genomes are a mosaic of modules with polyphyletic origins that can be linked to multiple unrelated viruses, either within and outside Psiviricota (Dolja, Kruovic and Koonin 2020, Gibbs et al. 2020). Despite their complex origin, emergence and diversification of modern potyvirids have been traced to plant-associated astro-like viruses (plastroviruses) and protopotyviruses, groups of viruses identified in plant transcriptomes and aquatic samples (Lauber et al. 2019, Wolf et al. 2020) (Fig. 1B).

Potyviridae includes > 200 plant virus species currently assigned to the twelve genera Arevirus, Bevemovirus, Brambyvirus, Bymovirus, Celavirus, Ipomovirus, Macluravirus, Poacevirus, Raymovirus, Rymovirus, Tritismovirus, and Potyvirus, with this last being the most speciose (Gibbs et al. 2020, International Committee on Taxonomy of Viruses 2020). Potyvirids have positive single-stranded RNA genomes of 8-11 kb that are translated into polyproteins, which are in turn hydrolyzed by viral proteinases to release a set of mature products (Adams, Antoniw and Beaudoin 2005, Rever et al. 2009, Garcia 2015). First studies of potyvirid genomes identified a basic layout with conserved gene abundance and order. Yet recent discoveries spurred by sequencing technological advances have revealed a large variability in the genomic structures and gene content. Potyvirid polyproteinases indeed show a common core led by diversified leaders that are enriched in non-core modules which expand the proteome structural and functional heterogeneity.

Here, we present a pan-family survey of the structural and functional diversity of the Potyviridae proteomes by delineating core and non-core modules (see Supporting Information). We pro-
provide a family-wide classification of P1 proteinases, and review knowledge of non-core domains. We examine abundance of the leader helper component proteinase (HC-pro) within the family, and uncover a putative papain-like protease domain in polyprotein leaders of known and putative Celavirus members. Using non-core module evolution as a case study, we summarize main molecular mechanisms that have acted in the Potyviridae radiation.

We also discuss the finding that common immune evasion roles can be identified in potyviral leader cistrons and those of plant, fungal and animal viruses; pointing to host adaptation as a main driver of their evolution. A perspective on the applications of potyviral leader proteinases and other non-core modules in biotechnology and synthetic biology is also presented.

Core and non-core modules of Potyviridae proteomes

Genera of Potyviridae have a common polyprotein core which is expanded by a heterogeneous array of non-core modules (Fig. 2). A set of eight mature proteins is conserved in the middle and carboxy (C) terminus of the polyproteins, namely P3, 6 kDa protein 1 (6K1), cytoplasmic inclusion (CI) protein, 6 kDa protein 2 (6K2), viral genome-linked protein (VPg), nuclear inclusion protein A proteinase (NIa-pro), nuclear inclusion protein B (NIb), and coat protein (CP) (Revers and García 2015). P3N-PIPO and P3N-ALT are generated by a frameshifting mechanism in the P3 cistron,
and are conserved (Yang, Li and Wang 2021, Choi et al. 2022). These conserved proteins have a common but polyphyletic origin (Gibbs et al. 2020). Nla-pro and Nlb are homologous to picorna-like signature genes, being, respectively, a chymotrypsin-like cysteine proteinase and an RNA-dependent RNA polymerase (RdRp) with phylogenetic affinity to animal-infecting Astroviridae and other Pisumviricota members. CP was likely acquired from other filamentous RNA viruses, whereas CI is a superfamily 2 helicase most closely related to flavivirid homologs (Koonin et al. 2008, Zamora et al. 2017, Dolja, Krupovic and Koonin 2020).

Organization of polyprotein amino (N) termini (leaders) is highly variable and bears distinctive genus- or even species-specific features (Fig. 2). Protein hidden Markov model (HMM) profiles allow for sensitive homology detection and have been applied to infer evolution of viral proteomes, as well as virus identification in metatranscriptomic datasets and taxonomic assignment (Nasir and Caetano-Anollés 2015, Wolf et al. 2018, Bin Jang et al. 2019, Callanan et al. 2020). A combination of HMM and protein profile scans was applied here to quantitatively survey the diversity and abundance of the Potyviridae non-core modules. P1 and HC-pro are the most common, but not universal, leader cistrons (Yang, Li and Wang 2021). Other non-core modules identified in few potyvirid species include the pretty interesting sweet potato potyviral ORF (PISPO), putative zinc fingers and DUF3725, alkylation B (AlkB), inosine triphosphate pyrophosphatase (ITPase/HAM1), as well as a superfamily 2 helicase most closely related to flavivirid homologs (Koonin et al. 2008, Zamora et al. 2017, Dolja, Krupovic and Koonin 2020).

Pretty interesting sweet potato potyviral ORF (PISPO)
The potyvirus sweet potato feathery mottle virus (SPFMV) has a large Type-A P1 and defective HC-pro (Yang, Li and Wang 2021). Transcriptional slippage takes place within P1 with the derived transcripts coding for a truncated P1 and the frameshift protein PISPO, which participates in RNA silencing suppression (Fig. 2, and see below) (Mingot et al. 2016, Untiveros et al. 2016). Besides SPFMV, PISPO is present in sweet potato virus 2, C, and G (Clark et al. 2012).

Zinc fingers and DUF3725
Zinc-finger domains mediate interaction with DNA, RNA, and proteins, and have a variety of cellular functions that include antiviral immunity regulation. A divergent cation coordinates two cysteines and histidines in CCHH zinc-fingers, but different cysteine/histidine compositions are found in the non-canonical CCHC, CCCH, and CCCC zinc-fingers, all of which have reported RNA interacting ability (Cassandri et al. 2017, Corley, Burnns and Yeo 2020, Wang and Zheng 2021).

Type-A P1s lack silencing suppression activity and biological roles of their putative zinc fingers remain unknown. The CCCC motif in some of them partially overlaps DUF3725 (Pfam: PF12523) (Fig. 4). DUF3725 is found in Streptomyces bacteriophage proteins (ATE85218.1) that share similarities to the zinc-binding domain of DnaG-like primases, which coordinates template binding and RNA primer synthesis in the replication of double-stranded DNA viruses (Gao et al. 2019).

Family-wide phylogenesis of the conserved protease domain supports the presence of two distinct lineages—Types A and B (Fig 3B). Type A is predominant (88%, Fig. 3C), and includes homologs that display plant co-factor dependency in vitro cleavage assays; it is found in all members of Potyvirus and Rymovirus, and in 3/7 of ipomoviruses (Fig 3D). Type-B proteinases do not need plant co-factors, displaying robust self-processing in multiple translation systems including bacteria (Rodamilans, Vali and Garcia 2013, Shan et al. 2018). This lineage is found in all Ipomovirus, Rymovirus, Poaceovirus, Triticovirus, and Brambyvirus members (Fig 3D). A tandem of both lineages is found in the ipomoviruses cucumber vein yellowing virus (CVYV), squash vein yellowing virus (SqVYV), and Cocccinia mottle virus (CocMoV) (Figs 2, 3B and D) Dombrovsky, Reingold and Antignus 2014, Dambiez et al. 2016).

Type-A P1 acts as a viral accessory factor, since deletion mutants are infectious and capable of replication and systemic movement (Rohožková and Navrátil 2011, Pasin, Simón-Mateo and García 2014). Consistent with its dispensability, ~10% of the recognized potyvirid species lack P1 (see Arenavirus, Bevemovirus, Bymovirus, Celavirus, and Macluravirus of Fig. 3D).

Diversity and evolution of non-core modules
P1 proteinases—two phylogenetically and biochemically distinct lineages
P1 is the least abundant among the potyvirid proteinases (Figs 2 and 3A). The C terminus includes a well-conserved chymotrypsin-like serine protease domain, a common module of RNA viruses, which autocatalytically releases P1 from the polyprotein (Rodamilans et al. 2018, Mann and Sanfaçon 2019). The N terminus is hypervariable, intrinsically disordered and dispensable for P1 proteolysis (Valli, López-Moya and García 2007, Pasin, Simón-Mateo and García 2014). It can tolerate sequence insertions and diverse atypical domains and functional motifs can be found within it (Fig. 2).

Potyvirus P1 is active in planta and in plant-based translation systems but its proteolysis is very low or absent in animal systems (Rohožková and Navrátil 2011). This supports the hypothesis that activation of potyviral P1 requires a plant co-factor, the identity of which is yet unknown.

Type-A P1s lack silencing suppression activity and biological roles of their putative zinc fingers remain unknown. The CCCC motif in some of them partially overlaps DUF3725 (Pfam: PF12523) (Fig. 4). DUF3725 is found in Streptomyces bacteriophage proteins (ATE85218.1) that share similarities to the zinc-binding domain of DnaG-like primases, which coordinates template binding and RNA primer synthesis in the replication of double-stranded DNA viruses (Gao et al. 2019).

Family-wide phylogenesis of the conserved protease domain supports the presence of two distinct lineages—Types A and B (Fig 3B). Type A is predominant (88%, Fig. 3C), and includes homologs that display plant co-factor dependency in vitro cleavage assays; it is found in all members of Potyvirus and Rymovirus, and in 3/7 of ipomoviruses (Fig 3D). Type-B proteinases do not need plant co-factors, displaying robust self-processing in multiple translation systems including bacteria (Rodamilans, Vali and Garcia 2013, Shan et al. 2018). This lineage is found in all Ipomovirus, Rymovirus, Poaceovirus, Triticovirus, and Brambyvirus members (Fig 3D). A tandem of both lineages is found in the ipomoviruses cucumber vein yellowing virus (CVYV), squash vein yellowing virus (SqVYV), and Cocccinia mottle virus (CocMoV) (Figs 2, 3B and D) Dombrovsky, Reingold and Antignus 2014, Dambiez et al. 2016).
characterization of AlkB roles in potyvirid infection remains to be addressed.

AlkB is embedded within replication-associated proteins of plant viruses in the families Alphalfexiviridae, Betaflexiviridae, Closteroviridae, and Secoviridae (Fig. 5). Phylogenomic analysis of plant viruses has highlighted a divergent evolutionary history for AlkB compared to other viral protein domains; it was concluded that AlkB probably emerged by multiple independent acquisition events (Bratlie and Drabløs 2005). For example, the divergent genomic organization and significant phylogenetic separation of BIVY and ENMV suggests that the two viruses acquired the domain independently (Fig. 5B). BIVY and ENMV have been identified in plants of the Rosaceae and Asteraceae, respectively (Susaimuthu et al. 2016), DNA bacteriophages (Yoshikawa et al. 2018), and giant DNA viruses (Fig. 5). A complex DNA methylation landscape was observed in genomes of the last of these (Jeudy et al. 2020), and viral AlkB may have roles in its regulation.

**Inosine triphosphate pyrophosphatase (ITPase/HAM1)**

ITPase is widespread in cellular organisms, hydrolyzing triphosphates of non-canonical purine nucleotides to prevent their incorporation in nucleic acids and preserve genome integrity (Simone, Pavlov and Borgstahl 2013).

A viral ITPase, also known as HAM1, was first identified in Ugandan cassava brown streak virus (UCBSV) and cassava brown streak virus (CBSV) (Figs 2 and 6). The two are atypical ipomoviruses that lack canonical Type-A P1 and HC-pro, encoding a single Type-B P1 with RNA silencing suppressor activity (Mbanzibwa et al. 2009, Dombrovsky, Reingold and Antignus 2014, Alicai et al. 2016, Shan et al. 2018). ITPase was later identified in *Euphorbia* ringspot virus (EuRSV, Potyvirus), encoding Type-A P1 and HC-pro (Fig. 6A) (Knierim, Menzel and Winter 2017). CBSV ITPase, although not essential for infection of experimental hosts, was involved in viral...
accumulation and symptom development. Contrary to the predicted antimutagenic activity of ITPass, viral mutation rates were not reduced in transgenic plants overexpressing CBSV ITPass, nor they were increased in CBSV clones lacking ITPass (Tomlinson et al. 2019). Use of improved sequencing approaches and alternative experimental systems could help shed light on the ITPass roles in potyvirid infection.

Metagenomics surveys have uncovered ITPass across diverse RNA and DNA virus taxa (Fig. 6). The ITPass fold is found in plant and invertebrate RNA viruses (Shi et al. 2016, Le Lay et al. 2020, and invertebrate RNA viruses (Shi et al. 2016, Le Lay et al. 2020, Anese et al. 2017, Gupta and Tatineni 2019a). For each virus, position of the first aligned polyprotein residue is indicated, and colored in red to label duplications; accession numbers are shown or given in Table S2 alongside virus complete names.

**Figure 4.** Putative zinc-finger motifs and DUF3725 in P1. Alignment of a conserved cysteine-rich region of Type-A and Type-B P1s is shown; inverted triangles indicate putative zinc-finger residues described to be involved in RNA silencing suppression activity of Type-B homologs (Vali, Dujovny and Garcia 2008, Kenesi et al. 2017, Gupta and Tatineni 2019a). For each virus, position of the first aligned polyprotein residue is indicated, and colored in red to label duplications; accession numbers are shown or given in Table S2 alongside virus complete names.
Leiva et al. 2022), as well as in bacteriophages and giant DNA viruses (Klijunen et al. 2005, Deeg, Chow and Suttle 2018, Sun and Ku 2021).

**Tobacco mosaic virus-like coat protein (TMV-like CP)**

Bymovirus is the only potyvirus genus whose member transmission is mediated by soil-borne plasmidophiroids (Jiang et al. 2020). Bymoviruses have bipartite genomes with RNA1 encoding the potyvirus polyprotein core, and RNA2, which encodes a second polyprotein processed in P2-1 and P2-2 (You and Shirako 2010). P2-1 is closely related to HC-pro (see below). P2-2 shares no similarity with other potyvirus proteins, and bymoviruses with its truncation or complete deletion are able to replicate and systemically move, but could not be transmitted by the natural vector (You and Shirako 2010).

Plasmidophiroid-transmitted viruses include *Virgaviroidae* and *Benyviridae* members (Tamada and Kondo 2013), whose capsid proteins show homology with bymovirus P2-2 (Dessens and Meyer 1996) (Fig. 7A). HMM-profile scans detect a ‘pseudo’ TMV-like CP domain conserved in P2-2 of all full-length bymovirus accessions and absent in oat mosaic virus, whose reference sequence is of a mechanically propagated isolate which lacks most of P2-2 (You and Shirako 2010). TMV-like CP sequences of bymoviruses cluster...
within a monophyletic clade, which supports their common origin (Fig. 7B). The P2-2 domains show phylogenetical relatedness to CP of Virgaviridae, and of wheat stripe mosaic virus (WhSMV), a putative benyvirus (Fig. 7B). Besides plant viruses, the TMV-like CP fold is found in algae and invertebrate RNA viruses but has negligible homology with cellular proteins (Nasir and Caetano-Anollés. 2015), and, based on current data, it can be safely con-
sidered a Potyvirus core component. Within a family-wide perspective, the reported genomic variation beyond the genus Potyvirus supports the HC-pro classification as a non-core module. Several ipomoviruses naturally lack HC-pro and its sequence is ab-
sent in ∼3% of potyvirid genomes (Figs 3A and 8A). Experimental evidence using a clone of wheat stripe mosaic virus (WSMV), a DNA virus, which encodes a defective HC-pro trans-complemented by a co-infesting potyvirus (Jayasinghe et al. 2021). Bymovirus RNA1 lacks leader proteases, whereas RNA2 encodes the HC-pro homolog P2-1 (Adams, Antonin and Beaudoin 2005), two function-
ally and phylogenically divergent HC-pro copies are present in arenaviruses (Qin et al. 2020) (Fig. 8A and B).

HC-pro has a papain-like cysteine protease domain that au-
tocatalytically hydrolyzes its C terminus (Guo, Lin and Ye 2011), and shows significant sequence divergence within the family that can be possibly rooted close to the beevomovirus ortholog and by-
movirus P2-1 (Fig. 8B). HC-pro shows homology to the nsP2 main protease of alphaviruses, as well as leader proteases of clo-
steroviruses, picornaviruses and arteriviruses (Gorbalenya, Koonen and Lai 1991, Mann and Sanfaçon 2019). Homology identification between HC-pro and Cryphonectria hypovirus 1 (CHV1) p29 was in-
strumental to postulate the evolutionary relationship between Po-
tyviridae and Hypoviridae, a family of fungal RNA viruses (Koonen et al. 1991) (Figs 1B and 8C). Papain-like cysteine proteases are common in cellular organisms, and main components of plant immunity (Misas-Villamil, van der Hoorn and Doehlemann 2016).

Helper component proteinase (HC-pro)

HC-pro is a multifunctional leader proteinase with roles in virus transmission, polyprotein processing, and suppression of antivi-
ral RNA silencing (Valli et al. 2018). The HC-pro RNA silencing sup-
pressor activity is indispensable for potyvirus infection (Garcia-
Ruiz et al. 2015), and, based on current data, it can be safely con-
sidered a Potyvirus core component. Within a family-wide perspective, the reported genomic variation beyond the genus Potyvirus supports the HC-pro classification as a non-core module. Several ipomoviruses naturally lack HC-pro and its sequence is ab-
sent in ∼3% of potyvirid genomes (Figs 3A and 8A). Experimental evidence using a clone of wheat stripe mosaic virus (WSMV), a DNA virus, which encodes a defective HC-pro trans-complemented by a co-infesting potyvirus (Jayasinghe et al. 2021). Bymovirus RNA1 lacks leader proteases, whereas RNA2 encodes the HC-pro homolog P2-1 (Adams, Antonin and Beaudoin 2005), two function-
ally and phylogenically divergent HC-pro copies are present in arenaviruses (Qin et al. 2020) (Fig. 8A and B).

HC-pro has a papain-like cysteine protease domain that au-
tocatalytically hydrolyzes its C terminus (Guo, Lin and Ye 2011), and shows significant sequence divergence within the family that can be possibly rooted close to the beevomovirus ortholog and by-
movirus P2-1 (Fig. 8B). HC-pro shows homology to the nsP2 main protease of alphaviruses, as well as leader proteases of clo-
steroviruses, picornaviruses and arteriviruses (Gorbalenya, Koonen and Lai 1991, Mann and Sanfaçon 2019). Homology identification between HC-pro and Cryphonectria hypovirus 1 (CHV1) p29 was in-
strumental to postulate the evolutionary relationship between Po-
tyviridae and Hypoviridae, a family of fungal RNA viruses (Koonen et al. 1991) (Figs 1B and 8C). Papain-like cysteine proteases are common in cellular organisms, and main components of plant immunity (Misas-Villamil, van der Hoorn and Doehlemann 2016).

Celevirion—an outlier

Celavirus is a single-member genus with celery latent virus (CelV) as the largest and most divergent of recognized potyvirids (Gibbs et al. 2020). CelV polypeptide initiates with a signal peptide that could translocate reporter proteins to the endoplasmic reticu-

Figure 6. HAM1/ITPase in Potyviridae and divergent virus taxa. (A) Diagrams of representative viral ORF including the inosine triphosphate pyrophosphatase (ITPase/HAM1) fold; relevant domains are colored. Left, taxonomic groups and species are shown; UCBSV, Ugandan cassava brown streak virus (Ipomovirus); EiRVS, Euphorbia ringspot virus (Potyvirus); E. coli ITPase rgdB is included as a standard. (B) Conserved residues in ITPase sequences. Alignment blocks show regions of E. coli rgdB that participate in substrate binding or catalysis (inverted triangle) (Savchenko et al. 2007). Position of the first residue is indicated (left), and (poly)protein size is shown in parentheses.
The identification of catalytic residues and cleavage sites in the CeLV leader nonetheless reveals the presence of a putative papain-like protease domain with sequence similarity to the HC-pro and bymovirus P2-1 domains. These domains are conserved in Striga potyvirus B (QKG60334; Fig. 8C), a virus phylogenetically related to CeLV and recently reported as Striga-associated potyvirus 2 (Choi et al. 2022).

Mechanisms of non-core module evolution

Our pan-family, quantitative survey of the Potyviridae proteomes defines the abundance of non-core modules and highlights discrete distribution patterns along the evolutionary tree of the family (Fig. 9A). High mutation rates, recombination, gene duplication and de novo emergence as well as extensive gene loss and gain, and host-niche adaptation drive virus evolution. Which are the main molecular mechanisms behind non-core proteome expansion in the family's evolutionary radiation?

Recombination is common in RNA viruses and an important component of potyvirus speciation (Sztuba-Solifaska et al. 2011). Its significance in potyvirus non-core module evolution can be clearly exemplified by AlkB identification in the potyvirus ENMV and the brambyvirus BlVY (Fig. 9B), possibly linked to independent acquisition events occurred in mixed infections with unrelated plant viruses.

Gene duplication is a major source of phenotypic novelty in cellular organism (Innan and Kondrashov 2010). It is however rare in RNA viruses, with the Closteroviridae coat protein duplication as a notable exception in plant viruses (Simon-Loriere and Holmes 2013). Tandem P1 or HC-pro copies in ipomoviruses and arenaviruses, respectively, were related to duplication events (Valli, López-Moya and García 2007, Qin et al. 2020). Empirical results nonetheless show that redundant sequences are rapidly purged from potyvirus genomes despite the potentially beneficial effect of the encoded proteins. Artificial insertion of a second HC-pro copy in the genome of TEV was deleterious and rapidly lost (Willemsen et al. 2016). Experimental evolution of a PPV clone encoding its own Type-A P1 and a second ortholog from a phylogenetically distant potyvirus led to an array of progeny viruses with enhanced fitness that were characterized by an almost or complete duplication loss (Rodamillas, Casillas and García 2021). Together the results indicate that both sequence identity and functional redundance constrain gene duplication in potyvirids. Further supported by the polyphyletic origin of duplicated copies (e.g. see Ipomovirus-encoded P1s labeled by II and V in Fig. 3), it can be concluded that gene duplication events detected in potyvirids are likely by-products of interspecific, ortholog recombination (Fig. 9B).

Neofunctionalization and functional specialization in potyvirids can be inferred from biochemical and biological characterization of P1 lineages (Fig. 9B). P1 was identified as a host adaptation determinant based on gene swapping and infection assays, and on genome-wide analysis of nucleotide variation (Salvador et al. 2008, Maliogka et al. 2012, Shan et al. 2015, Nigam et al. 2019), at a protein level, it shows conserved structural disorder (Pasin, Simón-Mateo and García 2014). Structurally flexible segments in viral proteins increase mutation tolerance and adaptability through acquisition of new linear motifs or protein domains (Gillin et al. 2014, Charon et al. 2018, Mishra et al. 2020). Strong evolvability and adaptation capacities of P1 are corroborated by family-wide identification of heterogenous motifs and domains within the P1 N termini, as well as the de novo emergence of PISPO through overprinting. Subfunctionalization allows the division of functions in duplicated genes (Innan and Kondrashov 2010). A zinc finger motif is conserved in all Type-B proteins but absent in most of P1s (Fig. 4); subfunctionalization of a Type-B-like ancestor could have participated in Type-A specialization.

Recently proposed scenarios place the bisegmented bymoviruses at the evolutionary diversification root of potyvird genera, which are suggested to have originated through genomic segment fusion (Qin et al. 2020). Celaviruses have monopartite genomes (Rose et al. 2019, Choi et al. 2022), and their ancestral status compared to bymoviruses as supported by RdRp phylogeny makes directed of the multipartite transition uncertain. Experimental examples are known of transitions from an originally non-segmented virus to a bisegmented one (Lucía-Sanz and Manrubia 2017). Supported by identification of a conserved pseudo TMV-like CP in bymoviruses, an intriguing possibility to explain the bipartidism emergence is the recruitment by a monopartite ancestor of a new genomic segment from co-infecting tobamoviruses to access to a vector transmission mode unprecedented within the family; transfer of the HC-pro homolog P2-1 could have been required to stabilize this de novo association and the gained multipartite state.
immune evasion through leader and non-core modules of potyvirids and other RNA viruses

In microbial systems, large taxonomic variability can be summarized by sets of redundant, polyphyletic functions (Louca et al. 2018). Functional analyses of cellular pangenomes suggest that gain of non-core genes influences adaptation of plant microbes to ecological niches (Box 1). Notwithstanding the low level of structural and biochemical conservation: Do Potyviridae non-core modules share biological function(s) and a common evolutionary driver? We present evidence supporting a main biological role of these non-core modules in counteracting host defensive reactions and thus host adaptation.

Potyvirid leader modules and RNA silencing evasion

Plant RNA viruses have evolved a variety of strategies to modulate disease severity and escape cellular antiviral responses (Paudel et al. 2018). In turnip mosaic virus (TuMV) HC-pro-proteolytic domains, the polyprotein sequence is split into two domains, the first with a dipeptide retention site and the second with a non-essential cleavage site. A previous analysis of the HC-pro domain of the Tobacco rattle virus (TRV) revealed a putative proteolytic domain (Figure S4). The validity of a common evolutionary driver of non-core modules is supported by the existence of a homologous proteolytic domain in the Celavirus genus, failed suppression by tobacco etch virus (TEV) protease, and the finding of the same proteolytic domain in the Striga potyvirus B (QVG60634.1) and Turnip mosaic virus (TuMV) HC-pro domains. The identity of the proteolytic domain in the Celavirus genus suggests a common evolutionary driver.
Figure 9. Potyviridae non-core proteome diversity and evolution drivers. (A) Abundance of non-core proteome components in genera of Potyviridae and recognized orphan species (CRCSV, SpMoV, LWBD); species numbers per genus are in parentheses; QVG60634.1, Striga potyvirus B; question marks indicate presence of putative homologs. Right, family phylogeny; RdRp domains were identified, protein sequences were aligned (Figure S5), and phylogeny was inferred; branches with bootstrap support ≥ 95 are in red, scale bar = 1. (B) Mechanisms and examples of non-core component evolution in Potyviridae. Non-core module inventory and virus complete names are given in Table S2.

and Sanfaçon 2018, Li and Wang 2019, Križnik, Baeble and Gru- den 2020). Being the first translation products, viral leader cistrons are considered important virulence and pathogenicity factors that can coordinate the early infection stages.

Potyvirid leaders are enriched in non-core modules (Fig. 2), and experimental evidence supports their roles in immune evasion and symptom development (Figs 10 and 11). RNA silencing is a major antiviral mechanism of plants. Potyvirus HC-pro is among the best characterized silencing suppressors, with multiple roles that include direct sequestration of small RNA molecules and inhibition of RNA silencing factors (Valli et al. 2018). In addition to HC-pro, other potyvirid leader proteins have been implicated in evasion of antiviral RNA silencing (Fig. 10A).
Type-A P1 does not suppress RNA silencing in transient expression assays, and enhances potyviral infection in plants defective in RNA silencing (Young et al. 2012, Pasin,Simón-Mateo and García2014), which suggests that the protein has additional roles independent of silencing suppression. P1 cis-expression strengthens the HC-pro activity, and improved translation in heterologous systems was implicated to this effect (Tena Fernández et al. 2013). P1 evolution for the mere enhancement of HC-pro expression appears unlikely, since optimization of the nucleotide Kozak context is not needed for silencing suppression of the ipomovirus SPMMV and the poacevirus Triticum mosaic virus (TriMV) (Giner et al. 2010, Gupta and Tatineni2019a). The P1 cistron was expressed alone, however, and it is unclear if proteolysis is needed during infection to release mature, active silencing suppressors from polyproteins.

GW motifs involved Type-B protein activity are also present in the potyviral PISPO, and the SPFMV P1N-PISPO fusion acts as a silencing suppressor that functionally replaces HC-pro (Mingot et al. 2016, Untiveros et al. 2016). GW motifs are present in HC-pro; although not involved in silencing suppression, they are needed to recruit AGO1 for pro-viral functions (Pollari et al. 2020), which further highlights the leader cistron multifunctionality in host adaptation.

**Negative autoregulation of potyvirid infection for immune evasion**

Diseases result from failures of cellular homeostasis (Kotas and Medzhitov2015). Negative feedback and incoherent feedforward loops are major autocontrol mechanisms that allow biological systems to adapt to changing environment and perturbations without homeostasis loss, disease or autoimmunity. They regulate natural and engineered cellular systems, as well as phase transitions and adaptation to resource changes of bacteriophages (Pitilli, Phukan and Coll2020, Brady et al. 2021, Frei and Khammash2021, Yao et al. 2021). The importance of negative autoregulation in plant virus infection just starts to be appreciated.
Mechanisms have been reported in plant viruses that avoid cellular toxicity or excessive inhibition of antiviral pathways, which can trigger host damage and pathogen fitness loss (Paulo and Sangaço 2018, Kriznik, Baebler and Gruden 2020). Promotion of RNA silencing spread was reported for tobamovirus movement protein and sobemovirus P1 (phylogenically unrelated to potyvirid P1), restriction of the silencing suppressor activity of cucumber mosaic virus 2b, geminiviral βC1 and polerovirus P0 was proposed to minimize host homeostasis perturbation (Vogler et al. 2008, Lacombe et al. 2010, Zhang et al. 2017, Ismayil et al. 2020, Watt et al. 2020, Clavel et al. 2021, Shukla et al. 2021).

HC-pro is a symptom determinant; its uncontrolled expression severely affects plant physiology, growth, fertility, and can trigger hypersensitive response or lethal necrosis (Pacheco et al. 2012, Valli et al. 2018). These findings suggest that tight control of proteins with strong silencing suppressor activity is desirable for optimal viral fitness, but how can it be achieved by an RNA virus that lacks transcriptional regulations? Recent data on P1 highlight a post-translational negative autoregulation that provides an evolutionary answer to the virus dilemma of countering defenses of the host without killing it. P1 can antagonize HC-pro, since the P1-HC-pro fusion lacks RNA silencing suppressor activity and could not sustain viral infection in hosts with unopposed antiviral immunity (Pasin, Simón-Mateo and García 2014). P1 is itself under autoinhibitory control. Several proteases display autoinhibitory domains or are synthesized as precursors that undergo structural rearrangements to activate (Hedstrom 2002, Gohara and Di Cera 2011, Trudeau et al. 2013). N-terminal deletions of PPV P1 identified a gain-of-function phenotype consistent with an autoinhibitory mechanism in which the N terminus negatively regulates P1 proteolysis, and self-cleavage results from autoinhibition relief by plant co-factor(s) (Pasin, Simón-Mateo and García 2014, Shan et al. 2018). A recent study model proposes the autoinhibited P1 self-cleavage as an immune evasion mechanism that regulates PPV replication through controlled release of the functional silencing suppressor HC-pro (Pasin et al. 2020). Self-controlled P1 processing kinetics would thus balance the strength of RNA silencing suppression with magnitude of phytohormone-mediated defense activation to mitigate resource burden and promote long-term viral fitness (Fig. 10B).

Additional immune evasion roles of potyvirid non-core modules

RNA silencing and other RNA metabolic pathways contribute to plant defense against potyvirids (Li and Wang 2019, Xu et al. 2020). They are further inter connected with autoimmunity, hormonal, and autophagic responses to provide robust plant immunity and tolerance to viruses (Cui et al. 2020, Pasin et al. 2020, Pitzalis et al. 2020, Shukla et al. 2021).

HC-pro interacts physically with RNA turnover components and inhibits EXORONUCLEASE 4 (XRN4) to counteract antiviral RNA decay (Fig. 10A) (Li and Wang 2018, De et al. 2020). Clover yellow vein virus (CICYVV) P1 was involved in overcoming the recessive resistance conferred by eu karyotic translation initiation factor 4E in pea (Nakahara et al. 2010). Selective translation enhancement of viral genomes by TEV P1 has been reported, which might contribute to suppressing expression of host immune factors (Martínez and Darbós 2014). An evasion strategy of al phavirus relies on disruption of stress granule formation by G3BP targeting mediated by peptide motifs that resemble the IxFG motif conserved in P1 N termini (Pasin, Simón-Mateo and García 2014, Panas et al. 2015, Reuper and Krenz 2021); P1 roles in stress granule processes are unknown.

Methylation impacts small RNA stability and loading in silencing complexes, and it is modulated by several silencing suppressors (Ji and Chen 2012, Csorba, Kontra and Burgvány 2015). HC-pro alters small RNA methylation through HUA ENHANCER1 methyltransferase interference and local disruption of the methionine cycle (Ji and Chen 2012, Ivanov et al. 2016, Del Toro et al. 2021). Roles of Alk8 and its RNA demethylase activity in potyvirid infection are less clear (Fig. 10A). It has been suggested that regulation of RNA methylation during infection contributes to viral immune evasion by fine-tuning viral replication rates or by post-transcriptional control of host gene expression (van den Born et al. 2008, Zhang, Qian and Jia 2021). N6-methyladenosine amount modulation was recently proposed as a new plant antiviral mechanism that hinders long-distance viral movement (Martínez-Pérez et al. 2021), and recent data indicate it could be effective against potyvirids as supported by reported changes in N6-methyladenosine levels upon bymovirus infection (Zhang et al. 2021).

HAM1 and its ITPase activity were recently shown to be CBSV necrosis determinants (Tomlinson et al. 2019). Although the mechanistic details were not studied, levels of inosine triphosphate (an ITPase substrate) were shown to regulate key factors potentially involved in antiviral immunity such as the viral RdRp catalytic ITPase substrate) were shown to regulate key factors potentially involved in antiviral immunity such as the viral RdRp catalytic speed and possibly viral replication rates, as well as plant stress response activation (Dulin et al. 2015, Kazibwe et al. 2020).

Recent results indicate possible P1 roles in coordinating plant homeostasis during mixed infections, since the protein impaired activity of the crinivirus silencing suppressor P25 (Domingo-Calap et al. 2021).

Expansion and immune evasion roles of RNA virus leaders

Diversification and immune evasion roles have been described for leaders or 5’ genomic cistrons of phylogenetically divergent groups of RNA viruses of plants, fungi and animals.

Plant RNA viruses of the Sobemoviridae family, genera Enamovirus and Polerovirus (family Luteoviridae), as well as Waikavirus and Fabavirus (family Secoviridae) belong to the picorna-like supergroup (Wolf et al. 2018). Cistrons encoded by their 5’ genomic portions show RNA silencing suppressor activity (Csorba, Kontra and Burgvány 2015, Sömera, Sarmiento and Truve 2015, Stewart et al. 2017, Carpino et al. 2020). Leader proteinases of Closteroviridae (phyllum Kentiviricota) affect pathogen virulence, superinfection exclusion, and promote viral amplification, possibly by viral replicase activation or subversion of host antiviral defenses (Dolja, Kreuze and Valkonen 2006, Atallah et al. 2016, Kang et al. 2018). Similar to potyvirids, proliferation of closterovirus leader proteinases is reported. A single, a tandem, or three copies of leader proteinases are found, respectively, in genomes of beet yellow virus, citrus tristeza virus, and actinidia virus 1, among others (Fig. 11).

Fungal RNA viruses of Hypoviridae recruit leader cistrons to counteract antiviral immunity. RNA silencing suppressor activity was reported for the CHV1 leader protease p29, and p24 of Cryphonectria hypovirus 4 (CHV4; Fig. 11) (Segers et al. 2006, Aulia et al. 2021).

Among animal viruses and similar to potyvirids, picornaviruses show expansion of genomic layouts with highly divergent leader (Fig. 11) (Gorbalenya and Lauber 2010, Zell 2018). Their leader proteins have a low level of structural and biochemical conservation, but share common biological functions in immune evasion.
Figure 11. Diversity and immune evasion roles of RNA virus leader cistrons. Genomes and encoded proteins of reference RNA viruses are represented as lines and arrowed boxes, respectively. Leader and 5' cistrons with immune evasion roles are colored and their names are on the right; proteinases are in boldface. Left, genera and species are indicated. Abbreviations and functional characterization references are as follows: PPV, plum pox virus; SPFMV, sweet potato feathery mottle virus; WSMV, wheat streak mosaic virus; TrMV, Triticum mosaic virus; CVYV, cucumber vein yellowing virus; UCBSV, Ugandan cassava brown streak virus; ANSSV, areca palm necrotic spindle–spot virus; CYNMV, Chinese yam necrotic mosaic virus; BIVY, blackberry virus Y (see main text). CHV1, Cryphonectria hypovirus 1; CHV4, Cryphonectria hypovirus 4 (Segers et al. 2006, Aulia et al. 2021). MCDV, maize chlorotic dwarf virus; BBVV1, broad bean wilt virus 1; FMDV, foot-and-mouth disease virus; EMCV-1, encephalomyocarditis virus 1; TMEV, Theiler’s murine encephalomyelitis virus; AnaV-A1, anativirus A1; EV7, echovirus 7 (Agol and Gmyl2010, Stewart et al. 2017, Freundt, Drappier and Michiels 2018, Lulla et al. 2019, Carpino et al. 2020, Saiz and Martinez-Salas 2021). CFMV, cocksfoot mottle virus; PEMV1, pea enation mosaic virus 1; PLRV, potato leafroll virus (Csorba, Kontra and Burgyán 2015). SHFV, simian hemorrhagic fever virus; PRRSV-2, porcine reproductive and respiratory syndrome virus 2; EAV, equine arteritis virus; PEDV, porcine epidemic diarrhea virus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2 (Han and Yoo 2014, Lunney et al. 2016, Shen et al. 2020, Nakagawa and Makino 2021). AcV1, actinidia virus 1; CTV, citrus tristeza virus; BYV, beet yellows virus (Dolja, Kreuze and Valkonen 2006). BVDV1, bovine viral diarrhea virus 1 (Tautz, Tews and Meyers 2015). Leaf icons indicate plant viruses; virus taxonomy and accession numbers can be found in Table S3.
(Agol and Gmyl 2010). Leader proteinase (LPc) of foot-and-mouth disease virus (FMDV, Aphthovirus) suppresses host cellular translation and antiviral responses by direct proteolysis of host translation factors and other RNA-binding proteins, signaling components, and conjugated ubiquitins (Saiz and Martinez-Salas 2021). Cardiovirus L, which is not a protease, antagonizes immune responses by suppressing interferon production, and can be functionally replaced by FMDV LPc (Freundt, Drappier and Michiels 2018, Visser et al. 2020). Theiler’s murine encephalomyelitis virus (TMEV, Cardiovirus) encodes the accessory L*, which directly targets RNase L ankyrin domains for interferon pathway inhibition and virus persistence promotion (Drappier et al. 2018). Murine but not human RNase L was found to be inhibited by L* (Drappier et al. 2018); this species-specific activity brings to mind the host-dependent activation of potyvirid Type-A P1 proteolysis (see above). The small protein UP was recently identified in the 5' region of diverse enteroviruses; UP modulates virus infection and tropism, and was suggested to participate in autophagy subversion for virus particle release (Lulla et al. 2019).

Leader size and domain organization vary considerably among members of the order Nidovirales. Arteriviruses are important veterinary disease agents; nsp1 is the first and most variable protein encoded. It is a leader proteinase, and up to three active copies are found in Deltaarterivirus (Vatter et al. 2014, Gulyaeva et al. 2017). Nsp1 proliferation resembles those of potyvirids encoding tandem of P1 or HC-pro (Fig. 2). Arterivirus nsp1 and its copies nsp1α, nsp1β, and nsp1γ counteract host immune defenses through interferon pathway suppression (Han and Yoo 2014, Lunney et al. 2016). Nsp1 of betacoronaviruses is released from polyprotein N termini to rapidly repress translation of cellular transcript and expression of innate immunity factors by 40S ribosomal subunit association (Nakagawa and Makino 2021). Immune evasion roles are conserved in nsp1 of alphacoronaviruses (Shen et al. 2020).

Pestivirus Npro (phylum Kitrinoviricota) is an accessory leader proteinase that acts as an interferon pathway antagonist to prevent cell apoptosis (Tautz, Tews and Meyers 2015, Jo et al. 2019).

Roles in immune evasion thus appear to be a functional link that connects Potyviridae non-core modules to each other, as well as leader cistrons of potyvirids with those of multiple RNA viruses (Fig. 11).

Biotech appeal of non-core modules

Infectious clones—established tools for potyvirid biological characterization and biotechnological advances

The accessory nature identified in non-core modules warrants the use of suitable experimental systems for their biological role characterization. Full-length infectious clones are universal, indispensable tools for virus biology research and the development of experimental systems for investigating diseases (Pasin, Menzel and Darös 2019, Kannan et al. 2020). They have been generated for members of Potyvirus (Domier et al. 1989), Tribetavirus (Choi et al. 1999), Macluravirus (Kondo and Fujita 2012), Foxtailcurvirus (Tatineni et al. 2015), Ipomovirus (Pasin et al. 2017), Celovirus (Rose et al. 2019), Areapovirus (Qin et al. 2020), as well as for bipartite viruses of Bymovirus (You and Shirako 2010, Ohki, Sasaya and Maoka 2019). Homology-based cloning methods are revolutionizing the potyvirid infectious clone construction, since they are efficient and require limited viral sequence information (Desbiez et al. 2012, Zhao et al. 2020). T-DNA vectors with stabilizing features have been used for one-step assembly of potyvirid clones suitable for Agrobacterium-mediated delivery (Pasin et al. 2017, 2018). A recently developed synthetic genomics framework with plant virome capacity could streamline characterization and engineering of plant viruses with no biological material need (Pasin 2021).

Non-core module characterization to guide plant expression vector development

Virus infectious clones can be engineered and optimized as expression vectors for plant biotechnology and synthetic biology (Fig. 12) (Pasin, Menzel and Darös 2019, Khakhar and Voytas 2021). Vectors based on potyvirids have been applied for disparate uses, ranging from production of heterologous peptides in plants, to flowering induction, gene silencing, metabolic engineering, CRISPR/Cas-targeted plant genome editing, and reprogramming of crops and their organelles (Lin et al. 2007, Llorente et al. 2020, Martí et al. 2020, Torti et al. 2021, Tuo et al. 2021, Uranga et al. 2021, Xie et al. 2021). Knowledge of Type-A P1 and its proteolytic activity has been instrumental in generating the first potyviral vectors (Fig. 12A). The bacterial β-glucuronidase (GUS) gene was inserted between TEV P1 and HC-pro, and the heterologous protein was released by polyprotein proteolysis mediated by P1 alone or in combination with Nla-pro (Dolja, McBride and Carrington 1992, Carrington et al. 1993). The same approach was used successfully in potyvirids encoding Type-B P1. GUS or fluorescent proteins were expressed using viral vectors derived from WSMV (Trirepartitevirus) (Choi et al. 2002, Tatineni et al. 2011), and TriMV (Tripartitevirus) (Choi et al. 2002, Tatineni et al. 2011), and TriMV
(Poaceivirus) (Tatineni et al. 2015). The 2A ‘self-cleaving’ peptides of FMDV or *Thiosea asigna* virus were applied to engineer Nla-pro independent processing of potyvirid polyproteins (Tatineni et al. 2011, Pasin, Simón-Mateo and García 2014).

The *ClYVV* P1/HC-pro junction was engineered for co-expression of multiple heterologous proteins that were released by P1 and Nla-pro proteolysis (Masuta et al. 2000). More recently, *ClYVV* was used for plant overexpression of a gibberellin catabolic enzyme inserted between P1 and HC-pro (Fig. 12B). Infections of pea and broad bean plants with the recombinant *ClYVV* conferred dwarfism, an agronomically important trait (Torti et al. 2021). Traditional plant breeding is time- and cost-consuming, and innovative strategies are needed for accelerated and tailored crop trait manipulation (Steinwand and Ronald 2020, French et al. 2021). Transient, viral-mediated manipulation of plant size and other agronomic performance traits holds promise to become a new standard for fast, flexible crop reprogramming.

The NIB/CP polyprotein junction is an insertion site used for heterologous gene expression that mimics the natural HAM1 location in poityvirids (Fig. 2). Simultaneous insertions at the P1/HC-pro and NIB/CP junctions allowed production of two recombinant proteins from a single potyviral vector (Beauchemin, Bougie and Kortemme 2020). *ClYVV* was used for plant overexpression of a gibberellin catabolic enzyme delivered to chloroplasts (Martí et al. 2020). In addition to protein overexpression, heterologous sequences inserted within P1 or at the P1/HC-pro junction can trigger silencing of plant homologs (Gammelgård, Mohan and Valkonen 2007, Xie et al. 2021). Potyvirid vectors have been used for virus-induced gene silencing, as well as for simultaneous plant gene silencing and heterologous protein production (Gammelgård, Mohan and Valkonen 2007, Tuo et al. 2021, Xie et al. 2021).

**Non-core modules—untapped synthetic biology resources**

Given their stringent specificity and orthogonality, potyvirid proteinases have been engineered for commercial purposes as well as for synthetic biology applications to control cellular functions (Chung and Lin 2020, Dyer and Weiss 2021). These proteinases have been integrated into synthetic signaling pathways with designs that included induction of degron-dependent protein depletion, autoinhibition release of transcription regulators, and enzyme reconstitution through dimerization inhibition or activation (Fernandez-Rodriguez and Voigt 2016, Gao et al. 2018, Fink et al. 2019). Use of potyvirid leader proteinases in synthetic genetic circuits has not yet been reported. Given its activation requirements and strict cis-cleavage activity, Type-A P1 could nonetheless be an appealing choice for biodesigns with high host specificity or bioc containment levels.

Synthetic, tight control over protein activity can be achieved by destabilizing tags, oligomerization domains, inhibitory modules, or subcellular sequestration signals (Alberstein, Guo and Kortemme 2021, Chen and Elowitz 2021). Type-A P1 was shown to undergo rapid degradation in plants and to inhibit activity of downstream fusion partners, such as HC-Pro or GUS (Verchot and Carrington 1995, Martínez and Daròs 2014, Pasin, Simón-Mateo and García 2014, Shan et al. 2015), and could be repurposed for conditional, fine-tuned activation of recombinant proteins.

HC-pro and other RNA silencing suppressors from plant viruses are used routinely to enhance protein yields of plant transient expression systems (Csorba, Kontra and Burgán 2015, Sainsbury 2020). Type-B proteins and P1N-PISPO could be also useful in similar applications. In-depth characterization of potyvirid AlkB and ITPase activities may also lead to novel tools for epigenetic or metabolic engineering applications.

**Leader proteinases—overlooked targets for antiviral strategies**

Human viruses are targeted by proteinase inhibitor therapies to a clinically useful level (Agbowuro et al. 2018); yet the use of similar antiviral strategies for plant virus control is lagging. Use of protease inhibitors for potyvirus control has shown limited success so far (Gutierrez-Campos et al. 1999), but new promising antiviral strategies have been reported. A plant protein involved in bacterial immunity was successfully repurposed to specifically sense Nla-pro and trigger antiviral cell death (Kim et al. 2016). This synthetic antiviral system has been implemented in soybean, and further optimized for enhanced control of the Nla-pro-induced cell necrosis (Helm et al. 2019, Pottinger et al. 2020). Potyvirid leader proteinases are attractive antiviral targets, since P1 and HC-pro defects preclude infectivity (Kasschau and Carrington 1995, Verchot and Carrington 1995, Pasin, Simón-Mateo and García 2014, Shan et al. 2015). A zucchini yellow mosaic virus isolate with reduced HC-pro silencing suppressor activity has been registered since 2007 for the US market as a cross-protection agent of cucurbits (U.S. Environmental Protection Agency 2007). RNA silencing transgenic approaches that target P1 or HC-pro confer potyvird resistance in crops (Di Nicola-Negri et al. 2005).

Investigation of non-core module roles in plant-potyvirid interactions recently allowed identification of new host factors and signaling pathways that could be exploited in antiviral strategies. High abscisic acid (ABA) levels were found to accumulate during infection of a FPV mutant having a truncated P1 (Fig. 10B), the finding prompted evaluation of ABA effects on infection. Defects of the cap-binding complex components HABA HYPERSENSITIVE1/CAP BINDING PROTEIN 80 (ABH1/CBP80) and CAP BINDING PROTEIN 20 (CBP20) are known to confer ABA hypersensitivity and were shown to significantly delay FPV infection (Pasin et al. 2020). Cap-binding complex contribution in antiviral defense was reported in other organisms, including insects and mammals (Gehhardt et al. 2019, Blagrove and Barrhieau 2021). ABA treatments promote resistance to FPV, and possibly to other potyvirids (Alazem, Widyasari and Kim 2019, Zhang et al. 2019, Pasin et al. 2020, Chiu et al. 2021). Rapid catabolism, photolability, and chemical instability make ABA unsuited for agricultural purposes.
Availability of synthetic ABA receptor agonists with high stability and binding affinities nonetheless paves the way for crop antiviral strategies based on chemical manipulation of ABA signaling (Hewage et al. 2020).

Research outlooks and conclusions

The phylum Pimoriviricota includes extremely diversified RNA viruses whose radiation was proposed to be concomitant with key eukaryogenesis events (Koonin et al. 2008). Potyviridae is currently the largest family of Riboviria (Fig. 1A), yet thousands of novel RNA viruses await accommodation in recognized taxa (Callanan et al. 2020, Edgar et al. 2022).

Gene gain and loss, specialization, and de novo emergence have promoted the diversification of leader layouts of Potyviridae (Figs 2 and 9), as well as of divergent RNA viruses of plants and animals, e.g. closteroviruses, picornaviruses and arteriviruses (Dolja, Kreuze and Valkonen 2006, Valli, López-Moya and García 2007, Agol and Gmyl 2010, Goryalena and Lauber 2010, Gulyaeva et al. 2017, Zell 2018). Functional expansion of a polyprotein core through domain gain is hypothesized to have taken part in the evolutionary transition from plastoviruses to modern potyvirids (Lauber et al. 2019). We point out that evolution of potyvirid non-core domains is diverse and can potentially be traced to multiple or single acquisition events (see AlkB in Potyvirus and Bramyvirus, or the pseudo TMV-like CP in Bymovirus, respectively; Figs 5 and 7), recombination and retention of functionally divergent homologs (P1 tandem in Ipomovirus; Fig. 3), as well as emergence of a new, overlapping protein module through overprinting (PIPO in Potyvirus). In-depth database search and sequence analyses uncovered a putative HC-pro-like domain within Celavirus (Fig. 8C), as well as the presence of the ITPase fold (HAM1, a former oddity of potyvirids) in taxonomically divergent RNA and DNA viruses (Fig. 6).

Identification of factors that interact with potyvirid non-core proteins and elucidation of host perturbations linked to their functional alteration are indeed major research priorities for dissecting their niche adaptation roles. Complete kinetic models were described for RNA viruses that share with potyvirids similar genome replication and protein expression strategies (Zitzmann et al. 2020, Lopacinski et al. 2021). Mathematical models could provide a quantitative understanding of the complex dynamics that regulate potyvirid replication as well as host immune responses and viral counterstrategies (Pasin et al. 2020). Accessory genes shape cellular pangenome diversity and are enriched in plant–microbe interaction determinants (Box 1), and models have been developed to describe pangenome gene content variation (Domingo-Sananes and McInerney 2021). Can empirical data from plant–potyvirid systems contribute to theoretical frameworks for understanding cellular pangenome evolution and ecological niche adaptation?

Functional redundancy of non-core modules has allowed the establishment of relationships between potyviral HC-pro and the non-canonical silencing suppressors P1N-PISPO and Type-B P1. HC-pro counteracts RNA decay antiviral defenses and associates with RNA turnover components for infection enhancement (Li and Wang 2018, De et al. 2020). Are these and additional HC-pro activities performed by other non-core modules? P1, HC-pro, and HAM1 are involved in symptom development (Valli et al. 2018, Tomlinson et al. 2019, Pasin et al. 2020). Is there any mechanistic connection between these otherwise structurally unrelated modules? Members of Nidovirales, the largest known RNA viruses, have evolved proofreading replication for maintaining integrity of genomes that can reach ~40 kb (Robson et al. 2020). Potyvirids are among the largest plant viruses and have unusually low mutation rates estimated to be in the range 10−5–10−6 mutations/site/generation (Sanjuán et al. 2010, Tomas and Elena 2010). Are functions carried out by AlkB, HAM1, or other modules conditioning the potyvirid evolution rates? Answers to these questions will assist in the better understanding contribution of the non-core proteome expansion in the Potyviridae evolutionary radiation and RNA virus evolution.

Finally, driven by advances in high-throughput sequencing technologies and easy access to underexplored geographical and ecological areas (Villamor et al. 2019, Maclot et al. 2020, Sommers et al. 2021), discovery of new potyvirids and poty-like ancestors with unusual genomic organization, atypical protein modules, and niche-optimized traits is likely to be further expanded in the near future (Lauber et al. 2019, Wolf et al. 2020). Extending efforts for potyvirid discovery and proteome functional characterization would improve understanding of non-core module roles in host adaptation evolution to eventually guide design of novel antiviral strategies and synthetic biology solutions.

Box 1. Non-core genes in plant–microbe interactions

Host niche adaptation is a major driving force of virus evolution (Simmonds, Aiewsakun and Katsourakis 2019), and functional characterization of the Potyviridae non-core modules supports their roles in symptom development and antiviral immunity evasion (main text). Pangenomes are increasingly used to represent known structural variants of cellular taxa, wherein the adaptive nature of accessory and rare genes is subject to debate (Domingo-Sananes and McInerney 2021, Coelho et al. 2022). Recent data nonetheless highlight the critical contribution of gene content variation in the evolution of plant-microbe dynamics and the genetic potential of holobionts (Badet and Croll 2020, Zilber-Rosenberg and Rosenberg 2021). Plant pathogenicity and adaptation factors are components of accessory genomes, lineage-specific replicons or chromosomes of bacterial and fungal species (Ma et al. 2010, Levy et al. 2017, Laflaame et al. 2020, Langner et al. 2021, Chou et al. 2022). Among plants, non-core genes of Arabidopsis thaliana, rice, rape, seed, cabbage, sunflower, and wheat are known actuators of the host–pathogen warfare (Zhao et al. 2018, Hübner et al. 2019, Van de Weyer et al. 2019, Bayer et al. 2020, Upadhyaya et al. 2021).

Supplementary data

Supplementary data are available at FEMSRE online.

Author contribution

FP conceived the work, performed data analyses, wrote the manuscript and prepared the figures; J-AD, and IET collaborated in the manuscript preparation. All authors revised and approved the final version.

Acknowledgements

The authors are grateful to AE Goryalena and AJ Gibbs for their comments, and to C Mark for editorial assistance. Computational resources from Galaxy are acknowledged.
Funding

FP is the recipient of a ‘Juan de la Cierva Incorporación’ contract [grant JC2019-039970-I] from Ministerio de Ciencia e Innovación (Spain). J-AD is funded by Ministerio de Ciencia e Innovación (Spain), and co-financed by European Regional Development Fund [grant PID2020-114691RB-I00]. IET is funded from crop commodity boards, the berry and ornamental industry, the United States Department of Agriculture and the National Science Foundation (USA).

Conflicts of interest.
None declared.

References

Adams MJ, Antoniw JF, Beaudoin F. Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. Mol Plant Pathol 2005; 6:471–87.

Agbouwo AA, Huston WM, Gamble AB et al. Proteases and protease inhibitors in infectious diseases. Med Res Rev 2018; 38:1295–331.

Agol VI, Gmyl AP. Viral security proteins: counteracting host defenses. Nat Rev Microbiol 2010; 8:867–78.

Alazem M, Widowarsi K, Kim K-H. An avirulent strain of soybean mosaic virus reverses the defensive effect of abscisic acid in a susceptible soybean cultivar. Viruses 2019; 11:887.

Alberstein RG, Guo AB, Kortemme T. Design principles of protein switches. Curr Opin Struct Biol 2022; 72:71–8.

Alicia T, Ndunguru J, Sseruwagi P et al. Cassava brown streak virus has a rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. Sci Rep 2016; 6:36164.

Atallah OH, Kang S-H, El-Mohtar CA et al. A 5’-proximal region of the Citrus tristeza virus genome encoding two leader proteases is involved in virus superinfection exclusion. Virology 2016; 489:108–15.

Aulia A, Hyodo K, Hisano S et al. Identification of an RNA silencing suppressor encoded by a symptomless fungal hypovirus, Cryphonectria hypovirus 4. Biology 2021; 10:100.

Badet T, Croll D. The rise and fall of genes: origins and functions of plant pathogen pathogenomes. Curr Opin Plant Biol 2020; 56:65–73.

Bayer PE, Golitz AA, Scheben A et al. Plant pan-genomes are the new reference. Nat Plants 2020; 6:914–20.

Beauchemin C, Bougie V, Laliberté J-F. Simultaneous production of plant pathogen pangenomes. BMC Biol 2022; 20:16.

Beauchemin C, Bougie V, Laliberté J-F. Simultaneous production of plant pathogen pangenomes. BMC Biol 2022; 20:16.

Beaupre S, Pollari M, Varjosalo M et al. Association of host protein VARICOSE with HCPro within a multiprotein complex is crucial for RNA silencing suppression, translation, encapsidation and systemic spread of potato virus A infection. PLoS Pathog 2020; 16:e1008956.

Begg CM, Chow C-ET, Suttle CA. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. eLife 2018; 7:e33014.

Bentley R, Sun H, Robinson C et al. In planta vs viral expression of HCPro affects its binding of non-plant 21–22 nt small RNAs, but not its preference for 5’-terminal adenines, or its effects
on small RNA methylation. New Phytol 2022;233;2266–81. DOI: 10.1111/nph.17935.

Desbiez C, Chandeysson C, Lecoq H et al. A simple, rapid and efficient way to obtain infectious clones of potyviruses. J Virol Methods 2012;183:94–7.

Desbiez C, Schoeny A, Maisonneuve B et al. Molecular and biological characterization of two potyviruses infecting lettuce in southeastern France. Plant Pathol 2017;66:970–9.

Desbiez C, Verdin E, Tepfer M et al. Characterization of a new cucurbit-infesting ipomovirus from Sudan. Arch Virol 2016;161:2913–5.

Dessens JT, Meyer M. Identification of structural similarities between putative transcription proteins of Polymyxa and Spongospora transmitted bymosviruses and furoviruses. Virus Genes 1996;12:95–9.

Di Nicola-Negri E, Brunetti A, Tavazza M et al. Hairpin RNA-mediated silencing of Plum pox virus P1 and HC-Pro genes for efficient and predictable resistance to the virus. Transgenic Res 2005;14:989–94.

Dolja VV, Kreuze JF, Valkonen JPT. Comparative and functional genomics of closteroviruses. Virus Res 2006;117:38–51.

Dolja VV, Krupovic M, Koonin EV. Deep roots and splendiferous boughs of the global plant virome. Annu Rev Phytopathol 2020;58:23–53.

Dolja VV, McBride HJ, Carrington JC. Tagging of plant potyvirus replication and movement by insertion of β-glucuronidase into the viral polypeptide. Proc Natl Acad Sci USA 1992;89:10208–12.

Dombrovsky A, Reingold V, Antignus Y. Ipomovirus – an atypical genus in the family Potyviridae transmitted by white flies. Pest Manage Sci 2014;70:1553–67.

Domier LL, Franklin KM, Hunt AG et al. Infectious in vitro transcripts from cloned cDNA of a potyvirus, tobacco vein mottling virus. Proc Natl Acad Sci USA 1989;86:3509–13.

Domingo-Calap MI, Chase O, Estapé M et al. The P1 protein of Watermelon mosaic virus compromises the activity as RNA silencing suppressor of the P2S protein of Cucurbit yellow stunting disorder virus. Front in Microbiol 2021;12:645530.

Domingo-Sananes MR, Mclnerney JO. Mechanisms that shape microbiobial pangenomes. Trends Microbiol 2021;29:493–503.

Drappier M, Jha BK, Stone S et al. A novel mechanism of RNase L inhibition: Thielers virus L protein prevents 2–5A from binding to RNase L. PLoS Pathog 2018;14:e1006989.

Dulín D, Vilfan ID, Berghuis BA et al. Elongation-competent pauses govern the fidelity of a viral RNA-dependent RNA polymerase. Cell Rep 2015;10:983–92.

Dyer RP, Weiss GA. Making the cut with protease engineering. Cell Chem Biol 2021;28451-9456(21)00515-8.

Edgar RC, Taylor J, Lin V et al. Petabase-scale sequence alignment catalyses virus discovery. Nature 602, 2022;142–7. DOI: 10.1038/s41586-021-04352-2.

Fedeleis BI, Singh V, Delaney JC et al. The AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases: repairing nucleic acid alkylation damage and beyond. J Biol Chem 2015;290:20734–42.

Fernandez-Rodriguez J, Voigt CA. Post-translational control of genetic circuits using Potyvirus proteases. Nucleic Acids Res 2016;44:6493–502.

Fink T, Lonzaric J, Praznik A et al. Design of fast proteolysis-based signaling and logic circuits in mammalian cells. Nat Chem Biol 2019;15:115–22.

Frei T, Khammash M. Adaptive circuits in synthetic biology. Curr Opin in Syst Biol 2021;28:100399.

French E, Kaplan I, Iyer-Pascuzzi A et al. Emerging strategies for precision microbiome management in diverse agroecosystems. Nat Plants 2021;7:256–67.

Freundt EC, Drappier M, Michiels T. Innate immune detection of cardioviruses and viral disruption of interferon signaling. Front Microbiol 2018;9:2448.

Gammelgård E, Mohan M, Valkonen JPT. Potyvirus-induced gene silencing: the dynamic process of systemic silencing and silencing suppression. J Gen Virol 2007;88:2337–46.

Gao XJ, Chong LS, Kim MS et al. Programmable protein circuits in living cells. Science 2018;361:1252–8.

Gao Y, Cui Y, Fox T et al. Structures and operating principles of the replisome. Science 2019;363:eaav7003.

Garcia-Ruiz H, Carbonell A, Hoyer JS et al. Roles and programming of Arabidopsis ARGONAUTE proteins during Turnip mosaic virus infection. PLoS Pathog 2015;11:e1004755.

Gebhardt A, Bergant V, Schnepf D et al. The alternative cap-binding complex is required for antiviral defense in vivo. PLoS Pathog 2019;15:e1008155.

Gibbs AJ, Hajizadeh M, Ohshima K et al. The potyviruses: an evolutionary synthesis is emerging. Viruses 2020;12:132.

Giner A, Lakatos L, Garcia-Chapa M et al. Viral protein inhibits RISC activity by Argonaute binding through conserved GW/GW motifs. PLoS Pathog 2010;6:e1000996.

Gitlin L, Hagai T, LaBarbera A et al. Rapid evolution of virus sequences in intrinsically disordered protein regions. PLoS Pathog 2014;10:e1004529.

Gohara DW, Di Cera E. Allostery in trypsin-like proteases suggests new therapeutic strategies. Trends Biotechnol 2011;29:577–85.

Gorbalenya AE, Koonin EV, Lai MM. Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. FEMS Lett 1991;288:201–5.

Gorbalenya AE, Lauber C. Origin and evolution of the Picornaeviridae proteome. In: Ehrenfeld E Domingo E Roos RP (eds.), The Picornaviruses. Washington, DC, USA: ASM Press, 2010, 253–70.

Gulyaeva A, Dunowska M, Hoogendoorn E et al. Domain organization and evolution of the highly divergent S’ coding region of genomes of arterviruses, including the novel possum nivovirus. J Virol 2019;91:e02096–16.

Guo B, Lin J, Ye K. Structure of the autocatalytic cysteine protease domain of potyvirus helper-component protease. J Biol Chem 2011;286:21937–43.

Gupta AK, Tatineni S. RNA silencing suppression mechanisms of Triticum mosaic virus P1: dsRNA binding property and mapping functional motifs. Viruses Res 2019a;269:197640.

Gupta AK, Tatineni S. Wheat streak mosaic virus P1 binds to dsRNAs without size and sequence specificity and a GWW motif is crucial for suppression of RNA silencing. Viruses 2019b;11:472.

Gutierrez-Campos R, Torres-Acosta JA, Sauco-Arias LJ et al. The use of cysteine protease inhibitors to engineer resistance against potyviruses in transgenic tobacco plants. Nat Biotechnol 1999;17:1223–6.

Han M, Yoo D. Modulation of innate immune signaling by non-structural protein 1 (nsP1) in the family Arteriviridae. Virus Res 2014;194:100–9.

Hedstrom L. Serine protease mechanism and specificity. Chem Rev 2002;102:4501–24.

Helm M, Qi M, Sarkar S et al. Engineering a decay substrate in soybean to enable recognition of the soybean mosaic virus NLa protease. Mol Plant Microbe Interact 2019;32:760–9.

Hewage KAH, Yang J-F, Wang D et al. Chemical manipulation of ascorbic acid signaling: a new approach to abiotic and biotic stress management in agriculture. Adv Sci 2020;7:2001265.
Holmes EC, Goldstein SA, Rasmussen AL et al. The origins of SARS-CoV-2: a critical review. Cell 2021;184:4848–56.

Hübner S, Bercovitch N, Todesco M et al. Sunflower pan-genome analysis shows that hybridization altered gene content and disease resistance. Nat Plants 2019;5:54–62.

Innan H, Kondrashov F. The evolution of gene duplications: classifying and distinguishing between models. Nat Rev Genet 2010;11:97–108.

International Committee on Taxonomy of Viruses 2020. ICTV master species list 2019 v1 (MSL #35). https://talk.ictvonline.org/files/master-species-lists/m/msl/9601 (June 2021, date last accessed).

Ismayil A, Yang M, Haxim Y et al. Cotton leaf curl Multan virus #C1 protein induces autophagy by disrupting the interaction of autophagy-related protein 3 with glyceraldehyde-3-phosphate dehydrogenases. Plant Cell 2020;32:1124–35.

Ivanov KI, Eskinel K, Bašić M et al. Molecular insights into the function of the viral RNA silencing suppressor HCPro. Plant J 2016;85:30–45.

Jayasinghe WH, Kim H, Sasaki J et al. Aphid transmissibility of onion yellow dwarf virus isolates with an N-terminal truncated HC-Pro is aided by leek yellow stripe virus. J Gen Plant Pathol 2021;87:178–83.

Jeudy S, Rigou S, Alempic J-M et al. The DNA methylation landscape of giant viruses. Nat Commun 2020;11:2657.

Je Li, Chen X. Regulation of small RNA stability: methylation and beyond. Cell Res 2012;22:624–36.

Jiang C, Kan J, Ordon F et al. Bymovirus-induced yellow mosaic diseases in barley and wheat: viruses, genetic resistances and functional aspects. Theor Appl Genet 2020;133:1623–40.

Jo WK, van Elk C, van de Bildt M et al. An evolutionary divergent pestivirus lacking the HEpg gene systemically infects a whale species. Emerg Microbes Infect 2019;8:1383–92.

Jones RAC, Naidu RA. Global dimensions of plant virus diseases: current status and future perspectives. Annu Rev Virol 2019;6:387–409.

Kang S-H, Atallah OO, Sun Y-D et al. Functional diversification upon leader protease domain duplication in the citrus tristeza virus genome: role of RNA sequences and the encoded proteins. Virology 2018;514:192–202.

Kannan M, Zainal Z, Ismail I et al. Application of reverse genetics in functional genomics of potyvirus. Viruses 2020;12:803.

Kasschau KD, Carrington JC. Requirement for HC-Pro processing during genome amplification of tobacco etch potyvirus. Virology 1995;209:268–73.

Kazibwe Z, Soto-Burgos J, MacIntosh GC et al. TORE mediates the autophagy response to altered nucleotide homeostasis in an RNAse mutant. J Exp Bot 2020;71:6907–20.

Kelloniemi J, Mákinen K, Valkonen JPT. Three heterologous proteins show that hybridization altered gene content and disease resistance. Nat Plants 2019;5:54–62.

Kong T, Fujita T. Complete nucleotide sequence and construction of an infectious clone of Chinese yam necrotic mosaic virus suggest that macluraviruses have the smallest genome among members of the family Potyviridae. Arch Virol 2012;157:2299–307.

Koonin EV, Choi GH, Nuss DL et al. Evidence for common ancestry of a chestnut blight hypovirulence-associated double-stranded RNA and a group of positive-strand RNA plant viruses. Proc Natl Acad Sci 1991;88:10647–51.

Koonin EV, Dolja VV, Krupovic M et al. Global organization and proposed megataxonomy of the virus world. Microbiol Mol Biol Rev 2020;84:e00061–19.

Koonin EV, Wolf YI, Nagasaki K et al. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nat Rev Microbiol 2008;6:925–39.

Kotas ME, Medzhitov R. Homeostasis, inflammation, and disease susceptibility. Cell 2015;160:816–27.

Križnik M, Babler Š, Gruden K. Roles of small RNAs in the establishment of tolerant interaction between plants and viruses. Curr Opin Virol 2020;42:25–31.

Lacombe S, Bangratz M, Vignols F et al. The rice yellow mottle virus P1 protein exhibits dual functions to suppress and activate gene silencing. Plant J 2010;61:371–82.

Laflamme B, Dillon MM, Martel A et al. The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. Science 2020;367:763–8.

Langner T, Harant, Gomez-Luciano LB et al. Genomic rearrangements generate hypervariable mini-chromosomes in host-specific isolates of the blast fungus. PLoS Genet 2021;17:e1009386.

Lauber C, Seifert M, Bartenschlager R et al. Discovery of highly divergent lineages of plant-associated astro-like viruses sheds light on the emergence of potyviruses. Virus Res 2019;260:38–48.

Le Lay C, Shi M, Buček A et al. Unmapped RNA virus diversity in termites and their symbionts. Viruses 2020;12:1145.

Lefevre P, Martin DP, Elena SF et al. Evolution and ecology of plant viruses. Nat Rev Microbiol 2019;17:632–44.

Leiva AM, Jimenez J, Sanival H et al. Complete genome sequence of a novel secovirus infecting cassava in the Americas. Arch Virol 2022;167:665–8.

Leyva A, Salas Gonzalez I, Mittel Viefhaus M et al. Genomic features of bacterial adaptation to plants. Nat Genet 2018;50:138–50.

Li F, Wang A. RNA decay is an antiviral defense in plants that is counteracted by viral RNA silencing suppressors. PLoS Pathog 2018;14:e1007228.

Li F, Wang A. RNA-targeted antiviral immunity: more than just RNA silencing. Trends Microbiol 2019;27:792–805.

Liang G, Bushman FD. The human virome: assembly, composition and host interactions. Nat Rev Microbiol 2021;19:514–27.

Lin M-K, Belanger H, Lee YJ et al. FLOWERING LOCUS T protein may act as the long-distance florigenic signal in the cucurbits. Plant Cell 2007;19:1488–506.

Llorente B, Torres-Montilla S, Morelli L et al. Synthetic conversion of leaf chloroplasts into carotenoid-rich plastids reveals mechanistic basis of natural chloroplast development. Proc Natl Acad Sci 2020;117:21796–803.

Lopacinski AB, Sweatt AJ, Smolko CM et al. Modeling the complete kinetics of cowpoxvirus B3 reveals human determinants of host-cell feedback. Cell Syst 2021;12:304–23.

Louca S, Polz MF, Mazel F et al. Function and functional redundancy in microbial systems. Nat Ecol Evol 2018;2:936–43.
Lucía-Sanz A, Manrubia S. Multipartite viruses: adaptive trick or evolutionary treat? *NFJ Syst Biol Appl* 2017;3:34.

Luilla V, Dinan AM, Hosmillo M et al. An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells. *Nat Microbiol* 2019;4:280–92.

Lunney JK, Fang Y, Ladinig A et al. Porcine reproductive and respiratory syndrome virus (PRRSV): pathogenesis and interaction with the immune system. *Annu Rev Anim Biosci* 2016;4:129–54.

Ma L-J, van der Does HC, Borkovich KA et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 2010;464:367–73.

Maclot F, Candresse T, Filloux D et al. Illuminating an ecological blackbox: using high throughput sequencing to characterize the plant virome across scales. *Front Microbiol* 2020;11:578064.

Majer E, Navarro J-A, Darós J-A. A potyvirus vector efficiently targets recombinant proteins to chloroplasts, mitochondria and nuclei in plant cells when expressed at the amino terminus of the polypeptide. *Biotechnol J* 2015;10:1792–802.

Maliogka VI, Salvador B, Carbonell A et al. Virus variants with differences in the P1 protein coexist in a *Plum pox virus* population and display particular host-dependent pathogenicity features. *Mol Plant Pathol* 2012;13:877–86.

Mann K, Sanfaçon H. Expanding repertoire of plant positive-strand RNA virus proteases. *Viruses* 2019;11:66.

Marti M, Diretto G, Aragonés V et al. Efficient production of saffron crocins and picrocrocin in *Nigellia sativa* using a virus-driven system. *Metab Eng* 2020;61:238–50.

Martínez F, Darós J-A. Tobacco etch virus protein P1 traffics to the nucleolus and associates with the host 60S ribosomal subunits during infection. *J Virol* 2014;88:10725–37.

Martínez-Pérez M, Aparicio F, López-Gresa MP et al. Arabidopsis m6A demethylase activity modulates viral infection of a plant virus and the m6A abundance in its genomic RNAs. *Proc Natl Acad Sci* 2017;114:10755–60.

Martínez-Pérez M, Gómez-Mena C, Alvarado-Marchena L et al. The m3A RNA demethylase ALKBHB plays a critical role for vascular movement of alfalfa mosaic virus in *Arabidopsis*. *Front Microbiol* 2021;12:745576.

Masuta C, Yamana T, Tacahashi Y et al. Development of clover yellow vein virus as an efficient, stable gene-expression system for legume species. *Plant J* 2000;23:539–46.

Mbanzibwa DR, Tian Y, Mukasa SB et al. Cassava brown streak virus (*Potyviridae*) encodes a putative Maf/HAM1 pyrophosphatase implicated in reduction of mutations and a P1 proteinase that suppresses RNA silencing but contains no HC-Pro. *J Virol* 2009;83:6934–40.

Mingot A, Valli A, Rodamilans B et al. The P1N-PISPO cassette delivers RNA virus infectious clones using the pLX mini binary vectors. *PLoS Pathog* 2020;16:e1004659.

Misas-Villamil JC, van der Hoorn RAL, Doehlemann G. Papain-like cysteine proteases as hubs in plant immunity. *New Phytol* 2016;212:902–7.

Mishra PM, Verma NC, Rao C et al. Intrinsically disordered proteins of viruses: involvement in the mechanism of cell regulation and pathogenesis. *Prog Mol Biol Transl Sci* 2020;204:1–78.

Morrison MS, Podracky CJ, Liu DR. The developing toolkit of continuous directed evolution. *Nat Chem Biol* 2020;16:610–9.

Nakagawa K, Makino S. Mechanisms of coronavirus nsp1-mediated control of host and viral gene expression. *Cells* 2021;10:300.

Nakahara KS, Shimada R, Choi S-H et al. Involvement of the P1 cistron in overcoming eIF4E-mediated recessive resistance against *Clover yellow vein virus* in pea. *Mol Plant Microbe Interact* 2010;23:1460–9.

Nasir A, Caetano-Anollés G. A phylogenomic data-driven exploration of viral origins and evolution. *Sci Adv* 2015;1:e1500527.

Nigam D, LaTourrette K, Souza PFN et al. Genome-wide variation in potyviruses. *Front Plant Sci* 2019;10:1439.

Ohta S, Takeda K, Maoka T. Cylindrical inclusion protein of wheat yellow mosaic virus is involved in differential infection of wheat cultivars. *Phytopathology* 2019;109:1475–80.

Pacheco R, García-Marcos A, Manzano A et al. Comparative analysis of transcriptomic and hormonal responses to compatible and incompatible plant-virus interactions that lead to cell death. *Mol Plant Microbe Interact* 2012;25:709–23.

Panas MD, Schulte T, Thaa B et al. Viral and cellular proteins containing FGDF motifs bind G3BP to block stress granule formation. *PLoS Pathog* 2015;11:e1004659.

Passin F, Bedoya LC, Bernabé-Orts JM et al. Multiple T-DNA delivery to plants using novel mini binary vectors with compatible replication origins. *ACS Synthetic Biology* 2017;6:1962–8.

Passin F, Menzel W, Darós J-A. Harnessed viruses in the age of metagenomics and synthetic biology: an update on infectious clone assembly and biotechnologies of plant viruses. *Plant Biotechnol J* 2019;17:1010–26.

Passin F, Shan H, García B et al. Abscisic acid connects phytohormone signaling with RNA metabolic pathways and promotes an anti-viral response that is evaded by a self-controlled RNA virus. *Plant Commun* 2020;1:100099.

Passin F, Simón-Mateo C, García JA. The hypervariable amino-terminal of P1 protease modulates potyviral replication and host defense responses. *PLoS Pathog* 2014;10:e1003985.

Passin F, Tseng X-A, Bedoya LC et al. Streamlined generation of plant virus infectious clones using the pLX mini binary vectors. *J Virol Methods* 2018;262:48–55.

Passin F. Oligonucleotide abundance biases aid design of a type IIS synthetic genomics framework with plant virome capacity. *Biotechnol J* 2021;16:2000354.

Paudel DB, Sanfaçon H. Exploring the diversity of mechanisms associated with plant tolerance to virus infection. *Front Plant Sci* 2018;9:1575.

Pitsilis E, Phukan UJ, Coll NS. Cell death in plant immunity. *Cold Spring Harb Perspect Biol* 2020;12:a036483.

Pitzalis N, Amari K, Grändorge S et al. Turnip mosaic virus in oilseed rape activates networks of sRNA-mediated interactions between viral and host genomes. *Commun Biol* 2020;3:702.

Polliari M, De S, Wang A et al. The potyviral silencing suppressor HCPro recruits and employs host ARGONAUTE1 in pro-viral functions. *PLoS Pathog* 2020;16:e1008965.

Pottinger SE, Bak A, Margets A et al. Optimizing the PBS1 decoy system to confer resistance to potyvirus infection in *Arabidopsis* and soybean. *Mol Plant Microbe Interact* 2020;33:932–44.

Qin L, Shen W, Tang Z et al. A newly identified virus in the family *Potyviridae* encodes two leader cysteine proteases in tandem that evolved contrasting RNA silencing suppression functions. *J Virol* 2020;95:e01414–20.

Rajamäki ML, Kelloniemi J, Alminaite A et al. A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions. *Virology* 2005;342:88–101.

Reuper H, Krenz B. Comparison of two Turnip mosaic virus P1 proteins in their ability to co-localize with the *Arabidopsis thaliana* G3BP-2 protein. *Virus Genes* 2021;57:233–7.

Revers F, García JA. Molecular biology of potyviruses. *Adv Virus Res* 2015;92:101–99.
Robson F, Khan KS, Le TK et al. Coronavirus RNA proofreading: molecular basis and therapeutic targeting. Mol Cell 2020;79:710–27.

Rodamilans B, Casillas A, García JA. P1 of Sweet potato feathery mottle virus shows strong adaptation capacity, replacing P1-HCPro in a chimeric Plum pox virus. J Virol 2021;95:e0015021.

Rodamilans B, Shan H, Pasin F et al. Plant viral proteases: beyond the role of peptide cutters. Front Plant Sci 2018;9:666.

Rodamilans B, Valli A, García JA. Mechanistic divergence between P1 proteases of the family Potyviridae. J Gen Virol 2013;94:1407–14.

Rohožková J, Navrátil M. P1 peptidase – a mysterious protein of family Potyviridae. J Biosci 2011;36:189–200.

Rose H, Döring I, Vetten H-J et al. Complete genome sequence and construction of an infectious full-length cDNA clone of celery latent virus – an unusual member of a putative new genus within the Potyviridae. J Gen Virol 2019;100:308–20.

Sainsbury F. Innovation in plant-based transient protein expression for infectious disease prevention and preparedness. Curr Opin Biotechnol 2020;61:110–5.

Saiz M, Martínez-Salas E. Uncovering targets of the leader protease: linking RNA-mediated pathways and antiviral defense. Wiley Interdiscip Rev RNA 2021:e1645.

Salvador B, Saénz P, Yanguez E et al. Host-specific effect of P1 exchange between two potyviruses. Mol Plant Pathol 2008;9:147–55.

Sánchez F, Ponz F. Presenting peptides at the surface of potyviruses in planta. In: Wege C, Lomonossoff GP (eds.), Virus-plant interactions and their roles in disease. New York, NY: Springer, 2018, 471–85.

Sanjuán R, Nebot MR, Chirico N et al. Viral mutation rates. J Virol 2010;84:9733–48.

Savary S, Willecoq L, Pethybridge SJ et al. The global burden of pathogens and pests on major food crops. Nat Ecol Evol 2019;3:430–9.

Savchenko A, Proudfoot M, Skarina T et al. Molecular basis of the antigenic activity of the house-cleaning inosine triphosphate pyrophosphatase RdgB from Escherichia coli. J Mol Biol 2007;374:1091–103.

Segers GC, van Wezel R, Zhang X et al. Hypovirus papain-like protease p29 suppresses RNA silencing in the natural fungal host and in a heterologous plant system. Eukaryotic Cell 2006;5:896–904.

Shan H, Pasin F, Tzanetakis IE et al. Truncation of a P1 leader protease facilitates potyvirus replication in a non-permissive host. Mol Plant Pathol 2018;19:1504–10.

Shan H, Pasin F, Valli A et al. The Potyviridae P1a leader protease contributes to host range specificity. Virology 2015;476:264–70.

Shen Z, Yang Y, Yang S et al. Structural and biological basis of alpha-coronavirus nsP1 associated with host proliferation and immune evasion. Viruses 2020;12:812.

Shi M, Lin X-D, Tian J-H et al. Redefining the invertebrate RNA virosphere. Nature 2016;540:539–43.

Shukla A, Hoffmann G, Kushwaha NK et al. Salicylic acid and the viral virulence factor 2b regulate the divergent roles of autophagy during cucumber mosaic virus infection. Autophagy 2021;1:1–3.

Simmonds P, Aiewsakun P, Katzourakis A. Prisoners of war - host adaptation and its constraints on virus evolution. Nat Rev Microbiol 2019;17:321–8.

Simon-Loriere E, Holmes EC. Gene duplication is infrequent in the recent evolutionary history of RNA viruses. Mol Biol Evol 2013;30:1263–9.

Simone PD, Pavlov YI, Borgstahl GEO. ITPA (inosine triphosphate pyrophosphatase): from surveillance of nucleotide pools to human disease and pharmacogenetics. Mutat Res Rev Mutat Res 2013;753:131–46.

Sõmera M, Sarmiento C, Truve E. Overview on sobemoviruses and a proposal for the creation of the family Sobemoviridae. Viruses 2015;7:3076–115.

Sommers P, Chatterjee A, Varsani A et al. Integrating viral metagenomics into an ecological framework. Annu Rev Virol 2021;8:133–58.

Steinwand MA, Ronald PC. Crop biotechnology and the future of food. Nat Food 2020;1:273–83.

Stenger DC, French R, Gildow FE. Complete deletion of Wheat streak mosaic virus HC-Pro: a null mutant is viable for systemic infection. J Virol 2005;79:12077–80.

Stewart LR, Jarugula S, Zhao Y et al. Identification of a maize chlorotic dwarf virus silence suppressor protein. Virology 2017;504:88–95.

Sun T-W, Ku C. Unraveling gene content variation across eukaryotic giant viruses based on network analyses and host associations. Virus Evol 2021;7:vpea081.

Susaimuthu J, Tzanetakis IE, Gengerich RC et al. A member of a new genus in the Potyviridae infects Rubus. Virus Res 2008;131:145–51.

Sztaba-Solíska J, Urbanowicz A, Figlerowicz M et al. RNA-RNA recombination in plant virus replication and evolution. Annu Rev Phytopathol 2011;49:415–43.

Tamada T, Kondo H. Biological and genetic diversity of plasmodesmoid-transmitted viruses and their vectors. J Gen Plant Pathol 2013;79:307–20.

Tatineni S, McMechan AJ, Bartels M et al. In vitro transcripts of wild-type and fluorescent protein-tagged Triticum mosaic virus (family Potyviridae) are biologically active in wheat. Phytopathology 2015;105:1496–505.

Tatineni S, McMechan AJ, Hein GI et al. Efficient and stable expression of GFP through Wheat streak mosaic virus-based vectors in cereal hosts using a range of cleavage sites: Formation of dense fluorescent aggregates for sensitive virus tracking. Virology 2011;410:268–81.

Tautz N, Tews BA, Meyers G. The molecular biology of pestiviruses. Adv Virus Res 2015;93:47–160.

Tena Fernández F, González I, Doblas P et al. The influence of cis-acting P1 protein and translational elements on the expression of Potato virus Y helper-component protease (HCPro) in heterologous systems and its suppression of silencing activity. Mol Plant Pathol 2013;14:530–41.

Tomlinson KR, Pablo-Rodriguez JL, Bunawan H et al. Cassava brown streak virus Ham1 protein hydrolyses mutagenic nucleotides and is a necrosis determinant. Mol Plant Pathol 2019;20:1080–92.

Torti S, Schlesier R, Thümmler A et al. Transient reprogramming of crop plants for agronomic performance. Nat Plants 2021;7:159–71.

Thomas N, Elena SF. The rate and spectrum of spontaneous mutations in a plant RNA virus. Genetics 2010;185:983–9.

Trudeau T, Nassar R, Cumberworth A et al. Structure and intrinsic disorder in protein autoinhibition. Structure 2013;21:332–41.

Tu D, Yan P, Zhao G et al. An efficient papaya leaf distortion mosaic potyvirus vector for induced gene silencing in papaya. Hortic Res 2021;8:144.

U.S. Environmental Protection Agency Zucchini yellow mosaic virus gene weak strain - chemical details. 2007. https://iaaspub.epa.gov/apex/pesticides/f?p=CHEMICALSEARCH:3:::1,3,31,7,12,25.P3_XCHEMICAL_ID:4462 (June 2021, date last accessed).

Untiveros M, Olspert A, Artola K et al. A novel sweet potato potyvirus open reading frame (ORF) is expressed via polymerase slippage and suppresses RNA silencing. Mol Plant Pathol 2016;17:1111–23.

Upadhyaya NM, Mago R, Fanvar V et al. Genomics accelerated isolation of a new stem rust avirulence gene-wheat resistance gene pair. Nat Plants 2021;7:1220–8.
Uranga M, Vazquez-Vilar M, Orzáez D et al. CRISPR-Cas12a genome editing at the whole-plant level using two compatible RNA virus vectors. CRISPR J 2021;4:761–9.

Valente JB, Pereira FS, Stempkowski LA et al. A novel putative member of the family Benyviridae is associated with soilborne wheat mosaic disease in Brazil. Plant Pathol 2019;68:588–600.

Valli A, Dujovny G, García JA. Protease activity, self interaction, and small interfering RNA binding of the silencing suppressor P1b from Cucumis melo yellowing (pumaviruses). J Virol 2008;82:974–86.

Valli A, López-Moya JJ, García JA. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. J Gen Virol 2007;88:1016–28.

Valli AA, Gallo A, Rodamilans B et al. The HCPro from the Potyviridae family: an enviable multitasking Helper Component that every virus would like to have. Mol Plant Pathol 2018;19:744–63.

Van de Weyer A-L, Monteiro F, Furzer OJ et al. A species-wide inventory of NLR genes and alleles in Arabidopsis thaliana. Cell 2019;178:1260–72.

van den Born E, Omelchenko MV, Bekkelund A et al. Tobacco mosaic virus in the host-virus interplay: Visser LJ, Aloise C, Swatek KN et al. Tobacco mosaic virus

Vogler H, Kwon M-O, Dang V et al. Tobacco mosaic virus movement protein enhances the spread of RNA silencing. PLoS Pathog 2008;4:e1000038.

Wang G, Zheng C. Zinc finger proteins in the host-virus interplay: multifaceted functions based on their nucleic acid-binding property. FEMS Microbiol Rev 2021;45:euaa059.

Watt LG, Crawshaw S, Rhee SY et al. The cucumber mosaic virus 1a protein regulates interactions between the 2b protein and ARGONAUTE 1 while maintaining the silencing suppressor activity of the 2b protein. PLoS Pathog 2020;16:e1009125.

Willemsen A, Zwart MP, Ambrós S et al. or not 2b: experimental evolution of functional exogenous sequences in a plant RNA virus. Genome Biol 2017;9:297–310.

Willemsen A, Zwart MP, Higueras P et al. Predicting the stability of homologous gene duplications in a plant RNA virus. Gen Biol Evol 2016;8:3065–82.

Willemsen A, Zwart MP. On the stability of sequences inserted into viral genomes. Virus Evol 2019;5:veoz045.

Wolf YI, Kazlauskas D, Iranzo J et al. Origins and evolution of the global RNA virome. Mol Bio 2018;9:e02329–18.

Wolf YI, Silas S, Wang Y et al. Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. Nat Microbiol 2020;5:1262–70.

Xie W, Marty DM, Xu J et al. Simultaneous gene expression and multigene silencing in Zea mays using maize dwarf mosaic virus. BMC Plant Biol 2021;21:208.

Xu M, Mazur MJ, Tao X et al. Cellular RNA hubs: friends and foes of plant viruses. Mol Plant Microbe Interact 2020;33:40–54.

Yang X, Li Y, Wang A. Research advances in potyviruses: from the laboratory bench to the field. Annu Rev Phytopathol 2021;59:1–29.

Yao T, Coleman S, Nguyen TVP et al. Bacteriophage self-counting in the presence of viral replication. Proc Natl Acad Sci 2021;118:e2104163118.

Yoshikawa G, Askora A, Blanc-Mathieu R et al. Xanthomonas citri jumbo phage XacN1 exhibits a wide host range and high complement of tRNA genes. Sci Rep 2018;8:4486.

You Y, Shirako Y. Bymovirus reverse genetics: requirements for RNA2-encoded proteins in systemic infection. Mol Plant Pathol 2010;11:383–94.

Young BA, Stenger DC, Qu F et al. Tritimovirus P1 functions as a suppressor of RNA silencing and an enhancer of disease symptoms. Virus Res 2012;163:672–7.

Yu B, Edstrom WC, Benach J et al. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. Nature 2006;439:879–84.

Zamora M, Méndez-López E, Agirrezabala X et al. Potyvirus viroid structure shows conserved protein fold and RNA binding site in ssRNA viruses. Sci Adv 2017;3:eaa02182.

Zell R. Picornaviridae – the ever-growing virus family. Arch Virol 2018;163:299–311.

Zhang T, Liu P, Zhong K et al. Wheat yellow mosaic virus NiB interacting with host light induced protein (LIP) facilitates its infection through perturbing the abscisic acid pathway in wheat. Biology 2019;8:80.

Zhang T, Wang Z, Hu H et al. Transcriptome-wide N6-methyladenosine (m6A) profiling of susceptible and resistant wheat varieties reveals the involvement of variety-specific m6A modification involved in virus-host interaction pathways. Front Microbiol 2021;12:656302.

Zhang W, Qian Y, Jia G. The detection and functions of RNA modification m6A based on m6A writers and erasers. J Biol Chem 2021;297:100973.

Zhang X-P, Liu D-S, Yan T et al. Metabolic evolution of functional exogenous sequences in a plant RNA virus. Genome Biol Evol 2017;9:297–310.

Zwillenberg D, Strecker I, Ruggieri A et al. Predicting the stability of homologous gene duplications in a plant RNA virus. Gen Biol Evol 2016;8:3065–82.

Zwillenberg D, Strecker I, Ruggieri A et al. Predicting the stability of homologous gene duplications in a plant RNA virus. Gen Biol Evol 2016;8:3065–82.