# Deep Roots and Splendid Boughs of the Global Plant Virome

Valerian Dolja, Mart Krupovic, Eugene Koonin

# Deep Roots and Splendid Boughs of the Global Plant Virome

Valerian V. Dolja[1], Mart Krupovic[2] and Eugene V. Koonin[3]

[1]Department of Botany and Plant Pathology and Center for Genome Research and Biocomputing, Oregon State University, Corvallis Oregon 97331-2902, USA; doljav@oregonstate.edu

[2]Institut Pasteur, Archaeal Virology Unit, Department of Microbiology, Paris, France; mart.krupovic@pasteur.fr

[3]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA; koonin@ncbi.nlm.nih.gov

## Key words:

plant virus, virus evolution, virus taxonomy, phylogeny, virome

## Abstract

Land plants host a vast and diverse virome that is dominated by RNA viruses, with major additional contributions from reverse-transcribing and single-stranded (ss) DNA viruses. Here we introduce the recently adopted comprehensive taxonomy of viruses based on phylogenomic analyses, as applied to the plant virome. We further trace the evolutionary ancestry of distinct plant virus lineages to primordial retroelements, RNA bacteriophages, circular bacterial ssDNA plasmids and, most notably, invertebrate RNA viruses. We discuss the growing evidence of the pivotal role that horizontal virus transfer from invertebrates and, to a lesser extent, fungi to plants played during the terrestrialization of these organisms. This process was enabled by the evolution of close ecological associations between arthropods, nematodes and fungi, on one hand, and plants, on the other hand. Ultimately, such associations that vary from predation and parasitism to symbiosis resulted in the evolution of the principal plant virus transmission routes that involve the respective vectors.

## I. Introduction: overview of virus origins, host ranges and megataxonomy

During the last decade or so, we have witnessed a dramatic inflation in the number and diversity of known viruses, thanks to the reduction of nucleotide sequencing costs and the rapid rise of metagenomics, metatranscriptomics and metaviromics. The preceding long-term stasis in the study of the global virome was marked with a heavy bias toward medically or economically important virus diseases. Although the remarkable diversity of virus genome replication and expression cycles was well appreciated, the virus world looked rather fragmented at the time. Somewhat paradoxically, the exponential growth in new virus discovery did not make the big picture even more patchy, but rather revealed numerous connections between virus lineages, enabled the development of several unifying concepts (12; 20; 21; 53; 68; 98; 144; 189; 195) and helped to draft the first coarse grain chart of the entire virus world (86). This brave new virus world emerged as a gene-genome network of high internal connectivity and perpetual dynamic exchange with the worlds of other mobile genetic elements (MGEs) and cellular host organisms (74; 83; 87).

Unlike cellular organisms that share ~100 homologous genes inherited from the Last Universal Cellular Ancestor (LUCA), there is not a single gene shared by all viruses. Therefore, as a whole, viruses are undoubtedly polyphyletic, originating on several or even numerous occasions from distinct gene sets (95). However, there are two functional types of virus genes that define the virus life style: those enabling semi-autonomous genome replication (replication modules) and those responsible for virion formation (morphogenetic modules). Some viruses encode no proteins directly involved in replication (67) whereas others have lost the morphogenetic module (i.e., capsid-less viruses) (83). However, the vast majority of viruses carry both replication and morphogenetic modules, and disentangling their evolutionary histories is key to understanding virus origins and evolution in general.

The semi-autonomous, semi-parasitic genome replication mode of viruses is shared with an enormous variety of selfish replicons or mobile genetic elements (MGEs) that do not form virions but encode at least some genes involved in their propagation either in host genome-integrated or extra-chromosomal state. The most common MGEs are DNA plasmids and self-propagating transposons (e.g., prokaryotic insertion sequences and

eukaryotic retrotransposons). Replication of the diverse MGEs involves protein-primed DNA polymerase B (PolB), rolling-circle replication endonucleases (RCRE), superfamily 3 helicase (S3H) or reverse transcriptase (RT). Strikingly, all these enzymes are also typical of viruses and are either rare or completely absent in cells (95). In addition, RNA viruses but not MGEs or cells encode RNA-dependent RNA polymerases (RdRPs) homologous to the RTs. It remains an open question whether the extant virus RdRPs are direct descendants of primordial RdRPs that might have been involved in the replication of RNA genomes after the evolution of translation in the hypothetical RNA world but before the advent of DNA genomes (189).

The shared evolutionary histories of replication modules of MGEs and viruses can be tentatively traced to ancient replication systems predating LUCA, with the primordial RNA recognition motif (RRM) domain being at the root of the replicative enzymes (95). Therefore, virus/MGE replication modules appear to emerge at the earliest stages of evolution, possibly even within precellular or protocellular replication systems, in accord with the concept of 'genetic parasite inevitability' (88).

What distinguishes viruses from other types of MGEs, are proteinaceous capsids that harbor and protect virus genomes between infections and enable genome delivery to the host cell. Despite the spectacular variety of capsid morphologies, the virus morphosphere is heavily dominated by icosahedral viruses, followed by those with elongated helical capsids. Recent analyses indicate that many if not most of virus morphogenetic modules have evolved from cellular ancestors at different phases of life evolution, from LUCA to this day (98). Thus, the evolutionary histories of the core modules of virus genomes point to ancient MGE-like elements providing replication-related proteins to emerging viruses while snatching protocapsid proteins from cells as the prevalent scenario of virus origins (95).

Given the apparent primordial origins of selfish replication modules, a timeframe for the origin of bona fide, encapsidated viruses can be approximated from the evolutionary history of the respective virus hosts. A paradigm of the early origins, perhaps at the LUCA stage, is provided by viruses with single and double jelly roll (SJR and DJR) CPs. These virion proteins were proposed to emerge on several occasions via repurposing of a wide variety of cellular carbohydrate-binding homologs sharing an SJR fold (98). Because the dsDNA viruses possessing icosahedral capsids formed by DJR-CP infect

bacteria, archaea and diverse eukaryotes, it seems likely that the DJR-CP evolved in the ancient, pre-LUCA virosphere, and viruses encoding this CP diversified adapting to newly evolving host organisms. Although SJR-CPs are particularly common in eukaryotic RNA and ssDNA viruses, there are some DNA bacteriophages and archaeal viruses utilizing this archetypic CP fold. However, it appears that these different groups of viruses have recruited cellular SJR proteins on several independent occasions.

In a sharp contrast to ancient and widespread SJR- and DJR-CPs, the RING domain virion matrix Z protein is utilized by only one family of vertebrate –RNA viruses, *Arenaviridae* (98). The Z protein fold is closely similar to that of eukaryotic E3 ubiquitin ligases implying relatively recent recruitment by ancestral arenavirus, apparently within a timeframe of vertebrate evolution. Therefore, emergence of viruses with novel combinations of replication and morphogenetic modules covers the entire history of life, from LUCA to vertebrates, and likely extends to this day (see section II c).

In general, viruses with distinct forms of encapsidated genomes (Baltimore classes) are differentially represented among evolutionary lineages of cellular host organisms (85). The archaea host the most restricted set of viruses, namely, dsDNA and ssDNA viruses only. In contrast, animals are the only group of organisms known to host viruses of all 7 Baltimore classes including RNA, reverse-transcribing and DNA viruses. The virome of the land plants has a distinct composition that is heavily dominated by diverse +RNA viruses, with a more limited representation of dsRNA, -RNA, reverse-transcribing and ssDNA viruses, to the exclusion of bona fide dsDNA viruses (33). In contrast, the virome of green algae is rich in large dsDNA viruses of the family *Phycodnaviridae,* apparently, at the expense of +RNA viruses (22; 124; 180). Furthermore, genomic remnants of distinct large dsDNA viruses have been identified integrated in moss genomes indicating that dsDNA viruses were banished from plants at a relatively late stage of evolution (119). The virome of fungi has a similar composition, lacking dsDNA viruses as well, but exhibits a bias toward dsRNA viruses (34; 46).

Although the Baltimore classes provide a useful framework for comparing virome compositions, the most recently developed and ICTV-approved classification of viruses is not based on Baltimore classes or virion morphology (or lack thereof in the case of capsid-less viruses) (85). Rather, this megataxonomy is underpinned by virus phylogenomics complemented by bi-partite (gene-genome) network analysis and

comparison of the virion and capsid protein structures. This evolutionary classification includes four virus realms, each subdivided into kingdoms, phyla, classes, orders, families, genera and virus species (85). The virome of land plants fits in two realms, *Riboviria* (RNA and reverse-transcribing viruses) and *Monodnaviria* (ssDNA viruses).

Below, we discuss the composition and large-scale evolution of the plant virome from the vantage point of phylogenomics. Because of the sparse sampling of 'lower' plants (green algae, bryophytes, lycophytes and ferns) as well as gymnosperms, the analysis of the plant virome is mostly limited to flowering land plants (angiosperms). However, viruses of lower plants are briefly covered in the context of the plant virome origins and evolutions. Our principal conclusion is that the plant virome was largely shaped by numerous events of horizontal virus transfer (HVT), often between extremely divergent hosts. The HVT events appear to occur through tight ecological association between diverse organisms including predation and parasitism as well as commensalism and symbiosis. The shorter-term evolution of viruses via mutations resulting in more limited host range expansion (45; 120) is beyond the scope of this article.

## II. Composition of the angiosperm virome

There are at least three metrics that are useful for classifying plant virome components: i) evolutionary, by phylogenomic and taxonomic diversity; ii) ecological, by the virus host range and infection frequency within plant populations; iii) economical, by virus disease impacts on crop, bioenergy or ornamental plants. Here we focus on the evolutionary approach, but also mention virus ecology and disease impacts where these are most relevant.

By and large, the replication and morphogenetic modules of plant viruses are shared with other viruses of eukaryotes, and animal viruses in particular (33). What distinguishes plant viruses from their kin, are processes defined by the specifics of plant biology: plant-to-plant virus transmission followed by two-phase systemic infection that involves local cell-to-cell movement and systemic transport through the plant vasculature (129).

The active intercellular virus spread occurs through plasmodesmata (14) and typically requires specialized, virus-encoded movement proteins (MPs) (62). The

evolutionarily diverse MPs represent the most prominent signature of plant viruses (126). Typically, systemic transport relies on additional, specific functionalities of different virus proteins such as MPs, CPs or counter-defense proteins (40). However, in some viruses with larger genomes, this function involves dedicated long-distance transport proteins.

The plant-to-plant transmission of viruses requires vectors such as plant-feeding arthropods, nematodes, plant-parasitic fungi and Plasmodiophorids (protists of the phylum *Cercozoa*) (9; 41; 135; 154). The process of transmission is virus- and vector-specific and often involves genetic determinants associated with virions and additional transmission factors known as 'helper components'.

Plants possess potent RNA-based defense systems against both RNA and DNA viruses including RNA interference (RNAi) also known as RNA silencing (6; 60; 107). To facilitate infection, many plant viruses rely either on specialized RNAi suppressor proteins or on suppression activity of proteins with other functionalities (e.g., MPs, CPs or transmission factors) (23).

Thus, a genome of an 'archetypal' plant virus contains replication, morphogenesis, transport, transmission and RNAi suppression modules. Many virus genes, particularly in viruses with small genomes, contribute to more than one of these activities. There are, however, plant viruses with reduced, minimal genomes that have either lost or have never acquired transport, transmission or counterdefense functions. Such viruses lead a persistent life style characterized by vertical transmission through seeds and/or pollen, and lack of pathogenicity and infectivity (ability to infect new hosts *de novo*, via horizontal plant-to-plant transmission) (137; 155). In this section, we describe the taxonomic structure of the global plant virome based primarily on the evolutionary provenance of the virus replication and morphogenetic modules.

a. RNA viruses: realm *Riboviria*, kingdom *Orthornavirae*

As mentioned above, most of the plant virome diversity fits into the realm *Riboviria*. Within this realm, the kingdom *Orthornavirae* harbors the bona fide RNA viruses with no DNA stage in their replication cycles (85). The replication modules of RNA viruses are organized around the RNA-dependent RNA polymerase (RdRP), the only gene that

is conserved in all viruses of this kingdom, to the exclusion of the rest of global virome. Therefore, the phylogenetic tree of the RdRPs is used as a scaffold to reconstruct the RNA virus evolution and to develop the corresponding taxonomy (189). According to this tree, *Orthornavirae* splits into 5 branches at the phylum rank (Fig. 1).

Phylum *Lenarviricota*. The deepest branching phylum *Lenarviricota* harbors +RNA bacteriophages that are believed to be the ancestors of eukaryotic virus families *Mitoviridae, Narnaviridae* and *Botourmiaviridae* (34; 189). The mitoviruses are capsidless RNA replicons that encode only the RdRP and replicate within the mitochondria. Most of the known mitoviruses have been identified in fungi, but recently, members of this family have been detected in plants as well (64; 138). Technically, capsid-less, non-infectious mitoviruses are mobile RNA elements, their claim to 'virusness' being solely the RdRP. The family *Botourmiaviridae* is also populated by fungal viruses, but contains a small genus *Ourmiavirus* that includes bona fide, encapsidated, MP-encoding plant viruses, of which the Ourmia melon virus discovered in Iran was the very first botourmiavirus (117; 150). In addition, a rich diversity of related, yet unclassified viruses has been described in invertebrates (168). Thus, plant ourmiaviruses represent but a twig within *Lenarviricota*, a phylum that is expected to spawn several new virus taxa.

Phylum *Pisuriviricota*. This second phylum of RNA viruses corresponds to a massive lineage previously described as 'picornavirus supergroup' (82; 89), and now splits into three classes, *Duplopviricetes, Pisoniviricetes* and *Stelpaviricetes* (Fig. 1) (85). The class *Duplopviricetes* consists of simple icosahedral dsRNA viruses that typically encode only the RdRP and a distinct type of capsid protein. There are two families including plant viruses in this class, *Partitiviridae* and *Amalgaviridae*. *Partitiviridae* is a vast family that includes a variety of fungal and unclassified invertebrate viruses (137; 168). The currently recognized plant partitivirids are corralled into two genera, *Alphapartitivirus* and *Betapartitvirus*, both shared with their fungal and unclassified invertebrate kin. This striking host range diversity among closely related viruses is suggestive of their exceptional propensity to HVT. A broad plant metavirome screening has shown that partitivirids are a prevalent component of the plant virome present in a variety of the

wild plant species (156). The reason why the apparent ecological dominance of partitivirids had been historically overlooked is that they lead a non-pathogenic life style. The dsRNA amalgavirids of plants and, again, fungi, encode RdRPs related to those of partitivirids and a protein distantly related to nucleocapsid proteins (NCs) of –RNA plant viruses in the *Phenuiviridae* family (see below) (94; 148). Similar to partitivirids, plant amalgavirids apparently resigned to a non-infectious, vertically transmitted life style (163). It is not clear if amalgaviruses form virions, but they provide an apparent case of a dsRNA virus that, instead of transcribing its genome *in virio*, which is typical of dsRNA viruses, adopted the replication mechanism involving NC that is characteristic of –RNA viruses.

In contrast to the dsRNA *Duplopiricetes* included into *Pisuriviricota* solely by virtues of the phylogenetic affinity of the RdRPs, *Pisonivirecetes* and *Stelpaviricetes* also share a chymotrypsin-like protease responsible for the polyprotein processing, the SJR CP and protein (VPg)-primed mechanism of RNA synthesis (Fig. 1). The class *Pisonivirecetes* includes two orders, *Picornavirales* and *Sobelivirales*, each harboring a family of icosahedral plant viruses, *Secoviridae* and *Solemoviridae*, respectively. The secovirids are rank-and-file picornaviruses with two-component genomes that share S3H with other *Picornavirales* (Fig. 1). Most of the family members are transmitted by insect (aphids, beetles, whiteflies) or by nematode vectors (164). By contrast, solemovirids have much smaller, densely-packed genomes (Fig. 1) and are transmitted by beetles and a variety of other insects (174).

Finally, the class *Stelpaviricetes* includes the order *Patatavirales* with a single, expansive, economically important plant virus family *Potyviridae* (47; 151). The potyvirids are a highly derived family of picorna-like viruses that encode a superfamily 2 helicase not found in other picorna-like viruses, as well as two additional proteases with multiple functions in virus replication, RNAi suppression and vector transmission. Unlike most of the viruses in this phylum that have icosahedral virions made of SJR CPs, potyviruses possess a distinct type of CP that forms flexuous filamentous virions of diverse plant +RNA viruses (fCP) (32; 193). Strikingly, structural analysis has shown that fCP is also homologous to phenuivirid (–RNA viruses) NCs, indicating yet another evolutionary connection between RNA viruses from different phyla (2). The potivirids in 9 of the 10 recognized genera share the non-propagative, non-persistent transmission

mode that involves virion attachment to receptors within the arthropod stylet or foregut mediated by the virus helper component (187). Interestingly, these potyvirus genera evolved affinity to distinct vectors including aphids, whiteflies and mites, whereas the tenth genus, *Bymovirus*, exploits Plasmodiophorid protists for virus transmission (154).

Phylum *Kitrinoviricota.* Unlike other *Riboviria* phyla where plant viruses are in the minority, the phylum *Kitrinoviricota* includes a large fraction of plant viruses which heavily dominate the class *Alsuviricetes* (85). Viruses in this class (formerly known as Alphavirus-like supergroup) share a universal signature of genome architecture that includes the capping enzyme (CapE), superfamily 1 helicase (S1H) and RdRP (Fig. 1)(82; 84). Aside from this three-component replication module, these viruses show remarkable diversity of genome organization and virion structure. On the minimalist end of the complexity spectrum, is family *Virgaviridae* that includes the archetypal tobacco mosaic virus (TMV) with a 6.4 kb genome which encodes only CapE-S1H-RdRP RNA replicase, MP and a single CP forming rigid rod-shaped particles (rCP). TMV employs an atypical, vector-less mode of transmission via mechanical damage of host plants, be it wind, passing animals or agricultural activities (165).

On the more baroque side is family *Closteroviridae,* where the prototype member, beet yellows virus (BYV), has a ~15.5 Kb genome encoding 10 proteins of which 5 form the morphogenetic module and assemble into complex filamentous virions (31). Three of these CPs are homologous to the fCP of potyvirids, as well as to those of alpha-, beta- and gammaflexivirids*,* also members of *Alsuviricetes* (32; 127). The six-component transport module of closteroviruses includes a dedicated MP and the entire morphogenetic module suggesting that a complex virion architecture evolved to facilitate virus movement (35). In addition, closteroviruses encode potent RNAi suppressors (19) and, altogether, present one of the most spectacular examples of genome complexification among RNA viruses (25; 35; 40). Analogous to potivirids, closterovirids from distinct genera are transmitted in a non-propagative, semi-persistent manner by different insect vectors, aphids, whiteflies or mealybugs (73).

Taxonomically, *Alsuvirecetes* split into three orders of which *Hepelivirales* harbors a single family of plant viruses, *Benyviridae* (49). The rCP of benyvirids is homologous to that of virgavirids (32) and forms the rod-shaped virions transmitted by

Plasmodiophorid vectors (154). The much larger order *Martellivirales*, in addition to *Virgaviridae* and *Closteroviridae* discussed above, includes *Bromoviridae*, *Kitaviridae* and *Endornaviridae*. The bromovirids have small tripartite genomes and icosahedral virions that are typically transmitted by various insects in a non-persistent manner. In particular, the most notorious of the bromovirids, cucumber mosaic virus, infects no less than 1,000 plant species and is transmitted by aphids (166). The kitavirids are the only plant viruses in this phylum with enveloped virions; these viruses are transmitted by mites (103; 145; 149) and are related to recently discovered insect viruses in the provisional genus "*Negevirus*" (181). The endornavirids are a peculiar group of viruses that, in addition to the replication module typical of *Alsuvirecetes*, encode various enzymatic domains, such as glycosyltransferase and capsular polysaccharide synthase, but have lost the morphogenetic module altogether. These capsid-less 'viruses' are found in fungi and oomycetes, but are extremely widespread in plants, where they cause symptomless, persistent, vertically transmitted infections that have been almost completely overlooked in the pre-metaviromics era (42; 158).

The third *Alsuviricetes* order, *Tymovirales*, consists of five families, including three families of filamentous viruses that share fCP, *Alpha-*, *Beta-,* and *Gammaflexiviridae*; (formerly, *Flexiviridae*), spherical *Tymoviridae* encoding SJR CPs and capsid-less *Deltaflexiviridae* (116). Among these, *Alpha-*, *Betaflexiviridae* and *Tymoviridae* infect plants, whereas *Gammaflexiviridae* and *Deltaflexiviridae* infect plant-pathogenic fungi. In addition to highly conserved RdRPs, CapE and S1H, many *Tymovirales* possess a papain-like protease. However, these plant virus families have unrelated transport modules: 'triple-gene block' MPs in *Alphaflexiviridae*, '30K-like' MP in *Betaflexiviridae* and a unique MP in *Tymoviridae* (126; 182). The alpha- and betaflexivirids are transmitted by a variety of arthropods, including aphids, mites and mealybugs, although viruses in the genus *Potexvirus* and some other genera appear to lack vectors and are transmitted mechanically (116). This latter property is apparently shared by some tymovirids, whereas others are transmitted by beetles in a non-propagative manner. Strikingly, it has been reported that tymovirids in the genus *Marafivirus* are transmitted by leafhoppers in a propagative fashion, that is, replicating within the insect (65), a feature that is so far unique among the non-enveloped +RNA plant viruses.

The second class in the phylum *Kitrinoviricota*, *Tolucaviricetes*, contains a single order *Tolivirales,* with two families of icosahedral plant viruses, *Tombusviridae* (186) and *Luteoviridae* (177). This class is linked to *Alsuviricetes* chiefly through the RdRP phylogeny; other genes encoded in the small, densely packed tombusvirus and luteovirus genomes encode only a SJR CPs (distantly related to those in tymoviruses) and unique types of MPs and RNAi suppressors (Fig. 1). Many of the tombusvirids are transmitted by fungi, whereas some (e.g., members of the genus *Tombusvirus*) could be soil-transmitted without vectors (154). The luteovirids are transmitted by aphids in a distinct, persistent, circulative, non-propagative manner, whereby viruses travel from the insect's gut through other tissues to salivary gland without replicating and are deposited through saliva when the aphid feeds on a next plant's phloem (9; 52).

The third class of *Kitrinoviricota*, *Flasuviricetes* (Flavivirus supergroup), currently includes the order *Amarillovirales* with a single family of exclusively animal, enveloped viruses, *Flaviviridae*. However, a single flavi-like virus, Gentian Kobu-sho-associated virus (GKaV), has been identified in gentians (alpine ornamentals) cultivated in Japan (4; 79). At ~23 Kb, this virus possesses the largest among all known plant viruses monopartite genome. Given the unusually high GKaV genome sequence variation within single infected plants, it seems possible that this unique flavi-like virus was relatively recently transferred from invertebrates to plants and is undergoing active adaptation to the new host. Indeed, the two viruses most closely related to GKaV are *Macrosiphum euphorbiae* virus 1 identified in a potato aphid (178) and Soybean cyst nematode virus 5 (7), both of these hosts being plant-feeding invertebrates.

Phylum *Duplornaviricota.* This phylum of dsRNA viruses encompasses a rather limited diversity of plant viruses that belong to two staggeringly dissimilar families, *Totiviridae* and *Reoviridae*. Totivirids are among the simplest RNA viruses encoding just a CP and the RdRP (Fig. 1), the same gene complement as in partitivirids. Furthermore, although the RdRPs of these dsRNA virus families are widely separated in the phylogenetic tree, both form similar icosahedral virions in which 60 homodimers of the CP are organized on a pseudo T = 2 lattice (118). This capsid architecture typical of diverse dsRNA viruses is not seen among other RNA or DNA viruses. The genome organizations of partitivirids and totivirids are, however, distinct: whereas the former

possess bi-partite genomes, totiviruses typically express CP and RdRP from a single genome-size mRNA via translational frameshift.

Plant-infecting totivirids have been discovered only recently in 'ecogenomics' studies. The persistent, apparently non-pathogenic life style of these viruses appears to be similar to that of plant partitivirids although much is to be learned on biology and ecology of plant totivirids (M. Roosinck, personal commun.). So far, *Totiviridae* has been known to include a variety of viruses infecting fungi, parasitic protists and invertebrates intermingled with the families *Chrysoviridae* and *Quadriviridae* that consist, mostly, of fungal viruses (46; 58; 168). Disentangling this phylogenomic and taxonomic quagmire awaits substantial effort that should be facilitated by further advances of metaviromics.

In contrast, plant reovirids are well-studied, pathogenic, insect-transmitted viruses (123; 184) that form three genera within the family that is otherwise heavily dominated by animal viruses and also includes a few fungal viruses and a virus from a green picoplankton alga (5; 168). Unlike the +RNA viruses discussed above, which are transmitted by arthropod or nematode vectors without replicating (non-propagative transmission), plant reoviruses replicate in their insect (leaf- or planthopper) vectors (propagative transmission). Such dual host range involving extremely divergent organisms provides a striking example of virus adaptability, as well as potential clues to the routes of reovirus evolution (see section IV below).

Similar to animal reoviruses, their plant kin possess large segmented genomes that consist of 10 or 12 unique dsRNA molecules (Fig. 1) encapsidated in a peculiar double-shelled, concentric icosahedral virion. These genomes endow reoviruses with coding capacity sufficient to produce up to 7 virion proteins including outer and inner CPs as well as RdRP and CapE that are co-encapsidated with the genome. These virion proteins function in virion assembly, vector transmission (outer CP) and RNA replication within infected cells. The non-structural proteins are involved in the formation of the viroplasm where virus reproduction and assembly apparently take place, RNAi suppression, as well as virus cell-to-cell movement in plants and insects (123).

Phylum *Negarnaviricota.* Similar to dsRNA viruses, –RNA viruses in this phylum encapsidate their RdRPs and additional replication proteins. Their virions, however,

adopt a highly distinct architecture, typically, with a condensed, helical, filamentous nucleocapsid and a membrane envelope adorned with virus-encoded glycoproteins (90).

The host range of *Negarnaviricota* is dominated by invertebrate viruses followed by vertebrate viruses (106; 167). Several –RNA viruses were recently discovered in protists (57) and fungi (108; 117). The plant viruses in this phylum are notably less diverse than their animal cousins, forming three families (*Tospoviridae, Fimoviridae* and *Aspiviridae*) and several genera within two large families of mostly animal viruses (*Phenuiviridae* and *Rhabdoviridae*). Again, similar to plant reovirids, most of these viruses are dual-host parasites that reproduce both in plants and in the arthropod vectors. Thus, plant *Tospoviridae* within *Bunyavirales* possess three-component ambisense genomes encapsidated into enveloped virions which also infect their minute insect transmission vectors, thrips (17). Perhaps, the most notorious of the plant –RNA viruses, Tomato spotted wilt virus, that is endowed with extremely broad host range and infects a variety of crops, is the prototype species in this family (1). The *Fimoviridae* of the same order are characterized by enveloped virions that harbor four- to eight-component genomes; fimovirids are transmitted in a propagative manner by mites, tiny arachnid arthropods (37). Finally, most plant rhabdovirids possess monopartite genomes typical of this family and are transmitted by and reproduce within the hemipteran insects including leaf- and planthoppers and aphids (*Cytorhabdovirus* and *Nucleorhabdovirus*) or arachnid mites (*Dichorhavirus*) (29; 188).

The –RNA plant viruses are closely related to their animal relatives in virion and genome architectures except for encoding MPs and RNAi suppressors that are required for systemic infectivity in plants (90). There are, however, three taxa of plant –RNA viruses that depart from this paradigm to different degrees. One such departure is seen in the genus *Tenuivirus* (*Phenuiviridae*) distinguished by segmented RNA genomes that are the largest among all known plant viruses (up to 8 segments totaling up to ~25 Kb) (104; 179). Apparently, the tenuiviruses have lost the ancestral membrane envelopes and switched to using their filamentous nucleocapsids as virions (38). However, tenuiviruses retained a non-structural glycoprotein as a helper component mediating the typical propagative mode of transmission by planthoppers (113). Likewise, the envelope-less, rod-shaped virions of *Varicosavirus* members (*Rhabdoviridae*) are formed by nucleocapsid-like CPs. The varicosaviruses exhibit a further notable

departure from the majority of arthropod-associated –RNA plant viruses in being transmitted by zoospores of soil fungi (188). In a similar manner, the aspivirids (previously classified as *Ophioviridae*) have also shed their envelopes and use nucleocapsids, albeit of unclear provenance, as fungus-transmitted virions (90; 154). Given the dominance of non-enveloped plant viruses, the evolution of these three virus taxa clearly reflects the adaptation of –RNA viruses to a more plant-specific life style.

b. Reverse-transcribing viruses

The second kingdom within the realm *Riboviria*, *Pararnavira*, consists of reverse-transcribing viruses encoding a reverse-transcriptase (RT) which is homologous to the RdRPs of RNA viruses. Hence, RNA viruses and reverse-transcribing viruses are assumed to have evolved from a common ancestor, warranting their classification within the same highest-level taxon (85). Among six officially recognized families within *Pararnavira*, plants host a share of *Metaviridae* and *Pseudoviridae* that encapsidate +RNA and entire family *Caulimoviridae* (informally referred to as pararetroviruses) that encapsidate dsDNA. The metavirids, pseudovirids and caulimovirids are classified into the phylum *Artvervicota*, class *Retraviricetes*, order *Ortervirales* (Fig. 1) (85; 93). All orterviruses share the replication module that consists of the RT and RNase H, a morphogenetic module (Gag polyprotein containing the characteristic α-helical CP and zinc-knuckle NC domains) and the polyprotein-processing aspartic protease (93; 97).

Similarly to vertebrate *Retroviridae*, metavirids and pseudovirids encode integrases and abundantly colonize cellular genomes across the eukaryotic branch of the tree of life, plants being no exception (111; 112). However, unlike infectious retrovirids, for most metavirids and pseudovirids, infectivity and intercellular spread have not been described, despite the conservation of the gene encoding structural Gag polyprotein and occasional presence (including in plant viruses) of genes encoding putative envelope proteins responsible for virus entry (102; 115; 191). Thus, metavirids and pseudovirids have been historically considered transposable elements and are more commonly known as long terminal repeat (LTR) retrotransposons of the Ty3/Gypsy and Ty1/Copia families, respectively (111). Accordingly, the majority of the identified LTR

retrotransposons have not been included into the ICTV framework, rendering the genus-level classification of these viruses incomplete (134).

Recent analysis of 80 plant genomes resulted in the identification of nearly 14,000 metavirids and pseudovirids (134). Both families are represented in all major groups of green plants, including the basal *Chlorophyta*, suggesting that both were present in the *Viridiplantae* genomes since their origin approximately 700–1500 million years ago (26; 112; 134; 143). Interestingly, neither plant metavirids nor pseudovirids encode recognizable MPs (134) suggesting that these viruses do not move between cells. Despite the lack of detectable infectious particles, widespread and frequent horizontal transfer of metavirids/pseudovirids in plants has been reported (36). Although the involved mechanisms remain unknown, high similarity between some fungal and plant metavirids suggests that plant-pathogenic fungi might participate in horizontal dissemination of metavirids and pseudovirids (142).

The caulimovirids share the replication and morphogenetic modules with metavirids and pseudovirids (Figure 1) but lead a radically different life style: they encapsidate circular dsDNA genomes and form infectious, isometric or bacilliform, non-enveloped virions (66). Phylogenetic analysis of the RT suggests that caulimovirids share the most recent common ancestor with metavirids (93). Similar to metavirids and pseudovirids, the basal caulimovirids, such as Petunia vein clearing virus, express all proteins as a single polyprotein, which is subsequently processed by the virus-encoded protease (153). Unlike the other orterviruses, replication of caulimovirids does not depend on integration into the host chromosome. Even though most caulimovirids do not encode an integrase, caulimovirus-derived endogenous virus elements (EVEs) are widespread in plant genomes and are thought to be integrated through non-homologous end-joining during DNA repair (16; 44). Although most of these EVEs are inactive, some have been shown to be fully infectious upon reactivation by various stress factors (43; 152).

Similar to many other plant viruses, caulimovirids encode a movement protein of the 30K superfamily (125) and are insect-transmissible in a genus-specific manner by aphids, mealybugs or leafhoppers. The mechanism of transmission involves two virus helper factors that bridge virions to the specific receptor at the tip of insect's stylet (10; 187). Notably, the reactivation-competent endogenous petuniavirus has no known insect vectors, further suggesting that it is a living intermediate between the metavirids and

more complex caulimovirids. Thus, evolution of caulimovirids from a metavirus-like ancestor likely involved the loss of the integrase gene and acquisition of the MP and vector transmission factors.

c. ssDNA viruses

Viruses with ssDNA genomes encoding rolling-circle replication endonucleases (RCRE) of the HUH superfamily are classified into the realm *Monodnaviria* currently encompassing six phyla (85). The phylum (unofficially referred to as CRESS-DNA viruses) includes all eukaryotic viruses with circular ssDNA genomes that encode homologous replicases (Reps) containing the N-terminal HUH endonuclease and C-terminal S3H domains (92; 197). This phylum unifies seven virus families and a vast number of viruses discovered by metagenomics that are affiliated tentatively (clades CRESSV1-6) (75). All plant viruses of this realm fall into two families (*Geminiviridae* and *Nanoviridae*) within the phylum *Cressdnaviricota* (Figure 2).

Most members of the phylum have some of the smallest genomes (~2 kb) found in the virus world and encode only 2 proteins, Rep and CP. Plant viruses in addition encode MP and RNAi suppressor. The virions of geminivirids have unique morphology of twinned (geminate) icosahedra encapsidating mono- or bipartite genomes (194). These virions are transmitted by insects (whiteflies, aphids, leafhoppers) in a genus-specific manner, in a circulative, non-propagative manner similar to that of +RNA luteovirids (187; 194).

By contrast, the genomes of nanovirids are partitioned into 6-8 circular DNA molecules of ~1 kb, each encoding a single protein and separately encapsidated into simple icosahedral virions (56). Remarkably, it has been demonstrated that different genomic segments of nanovirids rarely co-occur in the same cell; instead, they individually accumulate in distinct cells so that virus reproduction is achieved by complementation, whereby the gene products are shuttled between the cells (170). Furthermore, during circulative, non-propagative aphid transmission, the frequencies of genome fragment change, implying much more intimate relationships with the insect than simple passing tissue barriers from gut to salivary glands (171). Unlike

geminivirids, in addition to virions, nanovirid transmission requires a helper component (55).

Both geminivirids and nanovirids are associated with diverse satellite nucleic acids. For instance, viruses from both families support the replication of alphasatellites (*Alphasatellitidae*), which encode their own Reps but not the CPs, and thus depend on the helper viruses for transmission (13). The alphasatellites evolved from the Rep-encoding components of the nanovirus genomes, whereas the second genomic component of bipartite geminiviruses (DNA-B) could have originated from a satellite nucleic acid of unknown provenance (130). Moreover, the ultimate origin of eukaryotic CRESS-DNA viruses appears to be rooted in bacterial rolling-circle plasmids, and the diversity of these viruses apparently has been seeded on at least two independent occasions (74). Phylogenetic analysis indicates that the two classes of CRESS-DNA viruses in the phylum *Cressdnaviricota*, *Repensiviricetes* and *Arfiviricetes* that include geminivirids and nanovirids, respectively (Figure 2), evolved from two subgroups of related bacterial plasmids (74) implying that the two families are not monophyletic. The transformation of a plasmid into a virus, obviously, necessitated acquisition of a CP-encoding gene. Comparison of the CP structures has shown that CPs of different CRESS-DNA viruses have closer homologs among +RNA viruses than among other CRESS-DNA viruses. For instance, the CP of geminivirids is most closely related to that of satellite tobacco necrosis virus (63; 99), whereas CPs of cruciviruses are most similar to CPs of tombusvirids (28; 162). Thus, CRESS-DNA viruses appear to have evolved from plasmids through acquisition of reverse-transcribed CP genes (potentially, aided by RTs of endogenous reverse-transcribing viruses) from different groups of RNA viruses (74). In the case of plant ssDNA viruses, the CPs are at the forefront of virus interaction with the insect vectors. It has been suggested that adaptation to a new vector species could be more challenging than adaptation to a new plant host. Indeed, phylogenies of the geminivirid CPs mirror those of their vector species far more closely than those of the host species (105).

Remarkably, transformation of a plasmid into a virus is not a one-way street: geminivirids have apparently given rise to plasmids of phytoplasma (phloem-parasitic bacteria) by losing the CP gene (74). Evolution of the ssDNA viruses is thus one of the

most compelling manifestations of tight evolutionary connections between viruses and capsid-less MGEs, which appears to be a general trend in virus evolution (95).

## III. Plant evolution shapes the virome

The evolution of the 'green branch' of the eukaryotic tree of life (Archeoplastida) included successive emergence of the red algae (Rhodophytes), Glaucophytes, Chlorophyte and Streptophyte green algae, lineages that split ~1 billion years ago, and finally, a monophyletic lineage of the land plants (Embryophytes) that appeared ca. 400 mya (26; 27; 143). Apparently, the land plants evolved from the Zygnematophyceae branch of freshwater green algae which acquired resistance to desiccation and shared wet terrestrial habitats with the earliest Embryophytes, Bryophytes (hornworts, liverworts and mosses) (18). The evolutionary sequence within land plants branch following Bryophytes includes Lycophytes, Ferns, Gymnosperms (conifers), and finally, Angiosperms (flowering plants with two major lineages, monocots and eudicots). The Angiosperms diversified greatly in the Early Cretaceous, 140-100 mya, and flourished to dominate the terrestrial phytosphere as grasses, herbs, shrubs and trees (173). Virtually all agricultural output consists of the flowering plants, from rice to potatoes to apples to oranges, whereas timber production is based on both gymnosperm and angiosperm trees, and all of the flowering plants are hosting viruses.

What are the key evolutionary transitions in plant biology that are relevant to the formation of the contemporary virome of the land plants? One obvious consequence of terrestrialization is the switch from the aquatic to the soil/aerial life style. From the virus prospective, this life style change means losing the benefits of the aquatic environment that protects viruses outside the infected host from desiccation and UV damage as well as promotes virus dissemination via diffusion, convection and currents. By contrast, even in the wet soil environment, passive transmission of viruses between root systems of the host plants is very inefficient.

The land plant anatomy and cell architecture pose another set of limitations for plant-to-plant transmission of viruses, be it leaving the infected plant or entering a new one. The first barrier to virus penetration is epidermal cuticle, a layer of insoluble lipid polymers, such as polyester impregnated with hydrophobic waxes (139). There is simply

no way for a virus to penetrate undamaged cuticle except through open stomata formed by guard cells and functioning in gas exchange (51). Even if a virus manages to sneak through stomata into the leaf parenchyma, it faces the thick and rigid cell walls made of crystalline cellulose and matrix polysaccharides (e.g., hemicelluloses and pectins). These cell walls are an ancient feature shared by green plants: their composition is nearly identical in a lineage of Charophycean algae and land plants (147).

The multicellularity and complex vascular anatomy, which originated in land plants independently of other organisms (27; 140), pose additional severe challenges to viruses. For successful infection followed by plant-to-plant transmission, a virus must be able to move from cell to cell and/or through the vascular tissue, a route lying through plasmodesmata interconnecting plant cells and tissues. The Zygnematophyceae ancestors of land plants are unicellular or simple filamentous algae that lack plasmodesmata. Therefore, these essential organelles evolved *de novo* to mediate intercellular communications in land plants (14). Because the plasmodesmatal channels are narrow and highly-structured, they do not allow free passage of virions and serve as checkpoints for smaller macromolecular complexes.

Finally, both at the cellular and at the organismal levels, flowering plants possess potent innate immune responses to pathogens including viruses (15; 70; 107). The most powerful antiviral acquired immune response in land plants is RNAi (6; 60). In brief, RNAi is based on the recognition of 'abnormal' (highly structured and/or overexpressed) virus RNA, generation of small interfering RNAs (siRNAs) homologous to the virus genome and inactivation of this genome by the siRNA-guided Argonaute effector complex (39; 159). Importantly, induction of RNAi in a single virus-infected plant cell triggers amplification and systemic spread of the RNAi that follows or precedes the virus spread (121).

Although it is difficult to assess the exact contributions of soil/aerial life style, cuticle, cell wall, plasmodesmata and immunity to limitations in plant virome composition, one outcome of these defenses is a strictly reinforced taboo: no bona fide dsDNA viruses are allowed in land plants. This is in contrast to Chlorophyte algae where phycodnavirids with large dsDNA flourish (see below). Furthermore, integrated leftovers of distant phycodnavirid relatives are present in moss implying that these fossilized viruses have infected algal ancestors of moss (119). The reason for the

banishment of phycodnavirids in land plants could be that plasmodesmata are impenetrable for the large virions or dsDNA. Unlike phycodnavirids that break the cell walls via enzymatic digestion (172), none of the known viruses of land plants has this capacity.

However, what about small dsDNA viruses such as animal *Papillomaviridae* and *Polyomaviridae*? Both animal papillomavirids and plant reverse-transcribing caulimovirids encapsidate ~8 kb circular dsDNA genomes into 40-55 nm icosahedral virions. The caulimovirids move cell to cell through tubules formed by virus MPs implying that small, papillomavirid-like dsDNA viruses could have evolved a similar strategy. However, the host range of papillomavirids is limited to vertebrates, where they are highly host species-specific and tissue-restricted. Thus, potential explanation of their absence from plants could be the absence of a virus transfer route from vertebrates to plants.

The obvious follow-up question is: how have the extant, non-dsDNA viruses managed to prosper in land plants using their limited genomic resources? One fundamental solution is to surrender infectivity, that is, virus transmission between cells and plants, altogether. This solution is employed by 'cryptic' viruses that cause no disease and survive by means of vertical transmission through seed and pollen. This 'low-profile' life style is characteristic of minimalist persistent viruses that encode either RdRP alone (mitovirids) or RdRP and CP (partitivirids and totivirids) (137; 138; 155). The persistent endornaviruses have larger genomes of variable composition and appear to represent a transition state from the more aggressive life style of their ancestors in the *Alsuviricetes* lineage to persistence (42).

However, the majority of the known groups of plant viruses evolved a more radical strategy to cut through cuticle and cell walls, by hijacking plant-feeding organisms, such as invertebrates, fungi or protists, for vector-assisted penetration into and transmission between plant cells and tissues. In particular, 'piercing-sucking' insects deliver viruses by perforating the leaf cuticle and epidermal or phloem cell walls with their stylets (10; 135). Likewise, soil-dwelling ectoparasitic nematodes deliver stylet-borne viruses into root cells (41). The virus transfer by fungal or protist vectors into the root cells is achieved via encystment of zoospores that either absorb virus on their surface or internalize it (154).

As mentioned above, some viruses, for example, TMV eschew vector transmission and rely on stochastic mechanical transmission facilitated by the extreme environmental endurance that is characteristic of their virions. Another, relatively small subset of plant viruses take advantage of both horizontal (vectors or mechanical) and vertical (seed or pollen) transmission ensuring their long-term survival. This dual strategy is particularly important for viruses infecting annual hosts (61). Such viruses evolved means to cross the barrier between vegetative and reproductive tissues that protects plant progeny from the infection.

The strict requirements imposed by the multipronged host defenses on virus reproduction and transmission played a central role in shaping the composition of plant virus genomes and that of the global plant virome. Indeed, most of the functionalities of virus genes, beyond the basic requirements for genome replication and virion formation, are dedicated to virus-host interactions. These functions include cell-to-cell movement, long-distance transport, vector transmission, RNAi suppression and additional activities targeting immune responses. Of these, the dedicated MP and RNAi suppressor genes are most common and distinguish plant viruses from their kin infecting other eukaryotes. In the next section, we discuss the interplay between host and virus evolution that shaped the dynamic contemporary plant virome, as well as underlying evolutionary scenarios for major plant virus lineages.

## IV. Origins and diversification of the plant virome: horizontal virus transfer and virus-vector associations

Phylogenomics is the foundation on which the virome evolution concepts and scenarios rest. For *Viridiplantae*, our ability to reconstruct the path from viromes of the Rhodophytes, Glaucophytes, Chlorophytes, Streptophytes to those of Embryophytes in general and the Angiosperms in particular requires what is utterly missing – adequate sampling. We know precious little about viruses represented in most of these plant lineages (22; 124). The glowing exception is Angiosperms which are the basis of plant-based agriculture and thus are in the center of human attention. Hopefully, this state of affairs will change to the better before long: availability of hundreds of transcriptomes covering the entire plant kingdom (Archaeplastida) is a big step in this direction (143).

Despite these shortcomings, taking stock of the known viruses of alga is essential for our purpose. In Rhodophytes, the presence of dsRNA totivirid-like 'entities' was reported for two red macroalgal holobionts (100; 161); however, it is not clear if these viruses reproduce in algae or in associated fungi. Our literature search for viruses of Glaucophytes yielded no hits.

The current insight into the virome of *Chlorophyta* is more advanced, and so far, large dsDNA viruses of the family *Phycodnaviridae* steal the show, being found in a variety of marine picoplankton and freshwater algae (180; 185). In addition, Chlorophytes host small ssDNA viruses (8), a dsRNA reovirus (5), dsRNA partitivirus-like, capsid-less replicons (80; 81) and, potentially, a few other dsRNA viruses (124).

We know discouragingly little about the virome of *Streptophytae* (*Charophyta*) algae that include Zygnematophycean ancestors of land plants. Two very similar +RNA viruses related to benyvirids and virgavirids of the flowering plants have been identified in *Chara australis* and in fresh water metaviromes in Canada (48; 183). In addition, three RdRPs apparently belonging to dsRNA viruses have been detected in algal transcriptomes (124). Although deeper sampling of Zygnematophycean virome is much needed, it seems extremely unlikely that its diversity will ever approach that of the land plants. Indeed, there are fundamental differences in the biology and ecology of the complex vascular plants and comparatively simple freshwater algae that lack plasmodesmata and are not normally exposed to the most common vectors of land plants, flying arthropods. In addition, the species richness, a key determinant of the virome diversity, is ~100 times lower for *Zygnematophycea* compared to that of the vascular plants (Kew Botanical Gardens State of the World Plants, 2017; https://stateoftheworldsplants.org/2017/report/SOTWP_2017.pdf) (59). Therefore, it seems safe to assume that the algal ancestors of vascular plants could not harbor the seeds of the entire virus diversity represented in the extant land plants and rather served as a bottleneck in the virome evolution. Put another way, it appears likely that a substantial part if not most of the land plant virome was not inherited from the algal ancestors but was rather acquired via HVT from plant-associated organisms such as invertebrates, fungi and protists (34).

As radical as this claim might seem, it finds strong support in phylogenomic analysis of the rapidly growing data on the global virome, and in particular, RNA viruses that is

the dominant component of plant virome (85). One of the early realizations that plant viruses might originate from viruses of arthropods was concerned with –RNA viruses (33; 90). Most of these viruses have a dual host range, being arthropod-transmitted between plants in a propagative manner, that is, reproducing in both plants and vectors. Furthermore, the diversity of plant viruses is a subset of the arthropod virus diversity: in RdRp phylogenetic trees, plant-specific branches reside within a broader radiation of arthropod and arthropod/vertebrate viruses (106; 189). This same trend is prominent for both +RNA and dsRNA viruses in the phyla *Lenarviricota, Pisuviricota, Kitrinoviricota* and *Duplornaviricota* (168; 169; 189).

In addition to RdRP phylogeny, the invertebrate-to-plant HVT scenario for RNA viruses is supported by general evolutionary considerations. Despite a considerable margin of uncertainty, the current consensus is that plants started to colonize land somewhat earlier than invertebrates, but together they proceeded to form a terrestrial ecosystem after ~500 mya (77; 122; 160). At that time, invertebrates had already diversified greatly in the aquatic environments before and during Cambrian explosion (190). Accordingly, most of the currently known large-scale RNA virus diversity was likely present in aquatic invertebrates, such as mollusks and crustaceans (168), and followed invertebrates to land. In contrast, vascular plants just started to emerge then, going through their own Early Cretacean explosion when Angiosperms flourished merely 140-100 mya (173). Therefore, compared to flowering plants, invertebrates had a few hundred million years' head start to evolve their vast RNA virome, to bring it to land and to share it with evolving land plants.

Partial analysis of the primitive plant transcriptomes is also compatible with the above scenario by showing that the diversity of +RNA virus RdRPs grew along with land plant evolution form mosses to Angiosperms (124). Furthermore, virus MP gene transcripts were detected in lycophytes but not in the more ancient mosses, pointing to gradual virus adaptation to the growing plant complexity.

The composition of the extant biosphere is also in agreement with the evolutionary dominance of the invertebrate RNA virome the diversity of which is roughly proportional to the hosts' species richness. The terrestrial arthropods alone account for ~7,000,000 species, far exceeding all other land-dwelling eukaryotes combined (176), let alone vascular plants with only ~300,000 species (Kew Botanical Gardens State of

the World Plants, 2017; https://stateoftheworldsplants.org/2017/report/SOTWP_2017.pdf). Thus, plants are exposed to an enormous pool of invertebrate viruses which continuously sample the entire ecological space associated with the life styles of their hosts.

A survey of currently known plant virus vectors shows that insects, particularly *Hemiptera* (aphids, whiteflies, mealybugs, leafhoppers, planthoppers), *Thysanoptera* (thrips) and *Coleoptera* (beetles), hold the lead by transmitting +RNA, dsRNA, -RNA, ssDNA and pararetroviruses (187). In addition, mites of the *Arachnida* class of arthropods transmit some of the +RNA and –RNA viruses (29). Among the non-arthropod invertebrates, nematodes of the eponymous phylum *Nematoida* transmit +RNA secovirids (41).

As discussed above, the evolutionary scenario for the plant –RNA viruses seems to be the simplest: these viruses emerged from viruses of plant-feeding arthropods that acquired MPs and assorted RNAi suppressors, presumably, via recombination with pre-existing plant viruses (Fig. 3). Among these, plant tospovirids, fimovirids and rhabdovirids (*Cytorhabdovirus* and *Nucleorhabdovirus* genera) appear to be the least plant-specialized, possessing envelopes atypical for plant virome (17; 29; 37; 188). The tenuiviruses (*Phenuiviridae*) made a step toward 'plantness' by losing envelopes and repurposing a membrane glycoprotein as insect transmission factor (38; 113). Other envelope-less –RNA plant viruses, varicosaviruses (*Rhabdoviridae*) and aspivirids, are transmitted by soil-dwelling fungi raising the possibility of their origin via trans-kingdom HVT between fungi and plants (90).

The plant dsRNA reovirids that closely resemble their animal cousins (except for having acquired MPs and RNAi suppressors), reproduce in their Hemipteran vectors and lodge within clades of insect reovirids (123; 184), also fit insect-to-plant HVT scenario perfectly well. Among the +RNA viruses, kitavirids appear to follow the same paradigm, having kept their envelopes and the ability to reproduce in mite vectors, and being more similar to arthropod negeviruses than to any other plant viruses of the *Alsuviricetes* phylum (103; 181).

In contrast, most of the remaining plant +RNA viruses appear to be much more host-specialized, departing from their animal kin to different degrees. The point in case is *Alsuviricetes*, a class of the +RNA viruses that accounts for the largest share of plant

virome (Fig. 1). Indeed, 9 of the 15 officially recognized families in this class are plant-specific. Of these 9 families, five harbor viruses with the rod-shaped or filamentous virions formed by rCP and fCP, respectively, the two CP types historically believed to be plant virus-specific (32). That belief, however, predates the recent expansion of the known invertebrate virome, which resulted in the discovery of many viruses of this class in insects, chelicerates, myriapods, crustaceans, mollusks and nematodes. Phylogenetic analysis of RdRPs of this virus class placed plant-specific virus families, such as *Bromoviridae, Closteroviridae* and *Virgaviridae*, deep within the radiation of related invertebrate viruses, implying an ancestral relationship (168). Furthermore, the genomes of three invertebrate viruses (Behai charybdis crab virus 1 and Hubei virga-like viruses 2 and 9 from insects) contained the rCP-encoding genes (but no MP genes), supporting rCP origin within the invertebrate virome (168). The integrated copies of rCP genes were also 'excavated' from several fly genomes implying long-term presence of rCP-coding viruses in insects (78). Along the same lines, the evolutionary origins of the fCP were traced to the –RNA viruses (2).

Furthermore, within the large order *Tymovirales*, plant and insect viruses intermix in the *Maculavirus* genus (168). Even more 'damning' is a tymovirid genus of plant marafiviruses which are capable of reproducing in their insect vectors (65), potentially, an atavism going back to their arthropod ancestors. Collectively, phylogenomic analysis of the *Alsuvirecetes* once again points to evolutionary primacy of the invertebrate viruses over their plant 'offspring'.

This same trend of plant virus taxa 'nesting' within virus RdRP phylogenetic trees of animal/invertebrate viruses is apparent in the classes *Alassoviricetes* (plant *Ourmiavirus*), Pisonivirecetes (plant *Secoviridae* and *Solemoviridae*) (Fig. 3) and *Tolucaviricetes* (plant *Luteoviridae* and *Tombusviridae*) (Shi, Wolf)(168; 189).

A pertinent question concerning the non-propagatively transmitted viruses discussed above is how they acquire vector transmissibility upon switching to plant-only reproduction mode. In many of these viruses, the virion is the only essential transmission determinant that is apparently responsible for the receptor binding in the vector stylet or midgut and/or for guiding virus from arthropod's gut to salivary glands in the case of circulative transmission (9; 52; 135). Given the high evolvability of the virus CPs, adaptation to the receptor binding appears to be a relatively low evolutionary

barrier for viruses to cross. In addition, many plant viruses employ helper components/vector transmission factors, proteins specifically functioning in bridging virions to vector's receptors (10).

Although an invertebrate-to-plant HVT scenario seems to be prevalent in plant virome formation, distinct scenarios were also likely in action for several plant virus lineages. Thus, plant partitivirids in the *Duploviricetes* class share the family with fungal viruses implying possible HVT from fungi (157). The ancestry of class *Stelpaviricetes* that includes a single plant virus famil, *Potyviridae* is rather enigmatic having no close relatives among animal viruses except the phylogenetic affinity with the astrovirid RdRP. Apart from the RdRP, potyvirids and astrovirids share only homologous trypsin-like proteases that, however, do not appear to form a clade. In addition, the potyvirids encode fCP likely borrowed from other plant viruses, S2H related to those of flavivirids and several other proteins with unclear evolutionary provenance attesting to a highly mosaic origin of the potyvirid genomes (47). It seems likely that the ancestors of potyvirids are lurking somewhere waiting to be discovered by deeper metavirome sampling.

An intriguing nuance relevant to the enigma of potyvirid ancestry is offered by the viruses of the genus *Bymovirus* which, along with benyvirids (*Alsuviricetes*), use plasmodiophorid protists as vectors (154). Because these viruses and their proteins were found inside spores, they are likely capable of reproducing within the vector cells (72; 114). Remarkably, plasmodiophorids and related phagomyxids are cosmopolitan eukaryotic parasites of diatoms, oomycetes, brown algae and land plants that are prone to cross-kingdom shifts between these diverse hosts (133). Such unusual ecological mobility makes plasmodiophorids plausible vehicles of HVT from diverse aquatic protists to land plants, thus, short circuiting the need for invertebrate vectors. Indeed, diatoms and other protists are known to host a relatively diverse RNA virome that is considered ancestral to the vast RNA virome of invertebrates (34; 54).

Other striking departures from the invertebrate-to-plant HVT leitmotif are the evolutionary scenarios for ssDNA geminivirids and nanovirids, starting with two related but distinct bacterial plasmids and gradually evolving into plant-specific viruses (74). Their evolutionary paths included acquisition of distinct SJR-CPs from +RNA viruses followed by the virion adaptation for circulative, non-propagative insect transmission,

capture of the MPs from pre-existing plant viruses and, in the case of nanovirids, capture of the helper component from unknown source (Fig. 4).

Yet other evolutionary scenarios have been in action for the plant *Pararnavira*. In particular, the *Metaviridae* and *Pseudoviridae* are widely represented in diverse protists, fungi and plants (93; 111). The wide spread of mostly non-infectious metavirids and pseudovirids (better known as LTR retrotransposons) across eukaryotes implies their presence in the Last Eukaryotic Common Ancestor (LECA), but their subsequent history involved extensive horizontal transfer including between plants and fungi (36; 141). By contrast, *Caulimoviridae* are limited to plants, and *Retroviridae* to animals implying independent emergence from LTR retrotransposons (97). Thus, the evolution of the infectious life style of caulimovirids in plants has involved acquisition of MP and helper components (10; 66) but no interkingdom HVT.

In conclusion of this section, we need to consider the question of the ultimate origins of the genes defining the plant virus life style including MPs, vector transmission factors and RNAi suppressors. Although there are examples of clear appropriation of the host proteins for virus transport (e.g., Hsp70 homolog of closterovirids) as well as some cases of likely duplication of virus genes (e.g. the S1H of the triple gene block movement module) (35; 182), the majority of the diverse MPs including the ubiquitous TMV 30kDa-like MP have no detectable homologs outside plant viruses (126). Likewise, no direct ancestors were confidently identified for the virus transmission factors. It seems likely that these proteins evolved by exaptation of other host or virus proteins followed by rapid divergence erasing all traces of the ancestry. In the case of the MPs, the 'starting material' could be the host genes encoding proteins that possess cell-to-cell trafficking and RNA-binding capacities which are involved in plant development and anti-parasite defense (110). Similar to the extremely diverse MPs and transmission factors, virus RNAi suppressors are typically virus family-specific (23). Many of these proteins function by binding siRNAs and thus can be recruited from some of the numerous RNA-binding proteins available from host and virus genomes (19; 101). Alternatively, 'de novo' evolution, using recoding of pre-existing genes or chimeric genes arising through recombination is also a distinct possibility (30).

## V. Evolution of the overlapping RNA viromes of plants, fungi and animals

In this section, we present a brief overview of the relationships among the viromes of flowering plants, animals and fungi. The discovery of related RNA viruses in plants and animals was a veritable sensation in the early days of virus genomics (3; 50; 71). Since then, it has become clear that such relationships are a recurrent pattern in virus evolution that can be explained by HVT, by independent capture of homologous genes from hosts or other viruses, or by long-term coevolution with the hosts. The latter scenario implies a highly diverse virome in the LECA, given that plants and opisthokonts (animals and fungi) share common ancestry only at a very early stage of eukaryotic evolution (76). Distinguishing between these alternatives with confidence is difficult because the inadequate virome sampling complicates assessment of the virus spread across the host taxa. Nevertheless, for some groups of viruses, via combining the information on the depth of mixing in phylogenetic trees, evolutionary scenarios, and the biology of the host relationships, the most likely route of evolution becomes apparent.

The most pervasive phylogenetic blending of animal and plant viruses is observed among *Orthornavira* (Fig. 1). As emphasized in the previous section, RdRP phylogenies combined with genome architecture analysis for each of the five RNA virus phyla point to the most of the plant RNA virome evolving through HVT from much more diverse invertebrate virome (Fig. 5). This dominant scenario is strongly supported by tight biological associations between plants and invertebrates co-evolving along with land colonization.

A rather contrasting RNA virome evolution paradigm is apparent between invertebrates (broadly defined as all pre-chordate metazoa) and vertebrates. All major lineages of metazoa have diversified in Ediacaran era ~600 mya (146). Furthermore, the jawed vertebrates that comprise over 99% of modern vertebrates started to diversify at least 420 mya (11). Therefore, most of the major animal lineages were in place well before animal terrestrialization or diversification of the flowering plants. It is important to emphasize that all this animal diversity shared the marine habitat conducive to HVT. Accordingly, the early vertebrates were continuously sampled by invertebrate viruses forming the emerging vertebrate virome, and vice versa.

Although the invertebrate virome appears to be much larger than that of vertebrates, all major *Orthornavirae* lineages (except for *Lenarviricota*) are present in both invertebrates and early aquatic jawed vertebrates, fishes (167; 168). The viromes of the successive lineages of the terrestrial vertebrates, amphibians, reptiles, birds and mammals show strong signal of co-evolution with their respective hosts reflected in monophyletic, host-specific lineages in the RdRP phylogenetic trees (196). Thus, following plausible multidirectional HVT during diversification of animals in marine environment, as well as frequent, more recent HVT events (e.g., from blood-sucking arthropods transmitting arboviruses) (9; 196), the evolutionary scenario for vertebrates has a major virus-host co-evolution aspect (Fig. 5).

Enter fungi. The split between likely aquatic unicellular fungi and animals within opisthokont supergroup occurred very early in eukaryote evolution (76). However, unlike marine animals that diversified greatly before coming to land now ubiquitous mycelial fungi have flourished upon terrestrialization, reminiscent of the flowering plants (128; 175). According to the 'green scenario', land colonization involved association with the freshwater algal ancestors of land plants, accompanied by evolution of fungi decomposing organic matter and cycling nutrients, resulting in a tightly knitted 'phytomycobiome'. Therefore, evolutionary success of fungi and plants on land was mutually assured. In a course of terrestrial evolution, fungi diversified their life styles from plant-parasitic to being critical plant symbionts to saprophytic, gorging on plant remains. Upon colonizing land, fungi also evolved associations with other organisms including metazoa (128).

So far, the virome of the presumed ancient lineages of marine fungi remains largely unexplored, whereas in the terrestrial fungi, most of the available data deals with the plant-associated fungi (46; 117; 132). Despite these limitations, significant aspects of mycovirus diversity are apparent. One striking feature of fungal virome shared with the plant virome is complete absence of the dsDNA mycoviruses. The possible factors explaining this similarity include the fungal chitinous cell walls that block virus entry and the apparent lack of the virus-vectoring organisms that would surmount this barrier (34). Accordingly, majority of the mycoviruses possess no extracellular infectivity being transmitted vertically or through anastomosis (46). Another claim to originality is unusual richness of the fungal virome in mycovirus-dominated dsRNA virus families in

the classes *Chrymoviricetes* (totivirids, chrysovirids, quadrivirids, megabirnavirids) and *Duploviricetes* (partitivirids, amalgavirids). In addition, there is a propensity of fungi to hosting the capsid-less viruses (mitovirids, narnavirids, amalgavirids, hypovirids, deltaflexivirids) (46; 83).

Despite unexpected recent discoveries of the fungal –RNA mymonavirids (69) and unusual, extracellularly-transmissible ssDNA genomovirid (96; 192), the rest of the fungal virome appears to be borrowed from fungus-associated organisms, often plants. Indeed, many assorted mycoviruses and entire mycovirus families are derived from closely related plant viruses losing or repurposing their MPs and CPs in adaptation to fungal hosts (157). Thus, fungal *Deltaflexiviridae* and *Gammaflexiviridae* are derivatives of plant alpha- and betaflexivirids acquired via interkingdom HVT, whereas presumed hypovirid ancestor is related to potyvirids (24; 46).

The plant-to-fungus HVT is, however, bidirectional, with plant-specific mitovirids and partitivirids likely resulting from fungus-to-pant HVT (Fig. 5) (136; 137; 157). In the cases of botourmiavirids and endornavirids, the HVT direction is uncertain with present sampling. Many fungi are also associated with diverse animals (128) suggesting a strong potential for HVT between these organisms (91; 109; 131) (Fig. 5).

A hypothetical coarse-grain network of some major HVT and 'vertical' RNA virus-host co-evolution pathways is presented in Fig. 5. At the bottom, this network starts with bacteria that possess two RNA virus families, *Leviviridae* and *Cystoviridae* that spawned *Lenarviricota* and potentially, *Duplornaviricota* lineages of eukaryotic viruses, respectively. It has been also suggested that RTs from bacterial group II introns might have evolved into RdRPs of RNA viruses (189). The next step in eukaryotic RNA virus diversification has occurred in ancient protists, contemporary progenitors of which host relatively diverse RNA virome that, in turn, could have seeded explosive RNA virus diversification in the invertebrates (34). In addition to expanding then existing *Riboviria* lineages, the new lineages of *Negarnaviricota, Alsuviricetes, Flasuviricetes* and *Nidovirales* were apparently conceived at that time (Fig. 5).

Following its inflation, the invertebrate RNA virome served as a vast pool from which viromes of land plants have drawn generously. The virome of fungi, in addition to its ancestral dsRNA mycoviruses, has engaged in extensive two-way HVT with plants and likely animals (34; 157). Finally, the virome of jawed vertebrates has evolved via

both vertical and HVT-assisted virus acquisition from protists and invertebrates followed by lineage-specific virus-host co-evolution (196). However preliminary, this network provides a corner stone to a future complete picture of eukaryotic virus evolution achievable with comprehensive sampling of eukaryotes for viruses.

## VI. Concluding remarks

The plant virome covers two of the four realms of viruses, with the dramatic exception of the two dsDNA virus realms. Most likely, the exclusion of the dsDNA viruses that dominate the algal virome during the early evolution of plants is due to physical constraints, i.e. the inability of large and even moderate-sized dsDNA to pass through the plasmodesmata. The lack of dsDNA viruses in plants is compensated by the enormous diversification of +RNA viruses. Phylogenomic analysis of plant viruses demonstrates extensive phylogenetic mixing between viruses within numerous groups of RNA viruses as well as reverse-transcribing viruses. The plant virome appears to have been shaped by the interplay of four major evolutionary processes: 1) inheritance of a relatively small set of RNA and reverse-transcribing viruses from the algal ancestor, 2) acquisition of diverse viruses via HVT from invertebrates and fungi, 3) *de novo* emergence of ssDNA viruses (geminiviruses and nanoviruses) via recombination between plasmids and pre-existing RNA viruses, and 4) further, within plants diversification and adaptation of viruses acquired via each of the above three routes.

Although plant viruses have been classical objects of virology since its humble beginnings near the end of the 19$^{th}$ century, the investigation of the plant virome, until the last decade, was almost entirely limited to viruses that cause diseases in model and economically important plants. These studies have identified diverse groups of viruses but hardly could be expected to yield a comprehensive virome census. The advances of metaviromics in the last few years have changed this situation dramatically by expanding the virome and revealing the abundance of non-infectious, symptomless viruses, such as mitovirids, totivirids, partitivirids and endornavirids. This rapid progress in the characterization of the viromes of plants and other eukaryotes stimulated in depth phylogenomic studies which revealed the evolutionary trends outlined above and led to the creation of the all-encompassing megataxonomy of

viruses. Despite these major developments, we are far from the endgame because the current understanding of the viromes of plants and other groups of organisms is still based on sampling a small minority of the host diversity. However, the rapid advances of metaviromics suggest that representative sampling of the entire earth virome could be achievable within one to two decades. At that stage, comprehensive phylogenomic analysis (obviously, a challenge in its own right) will show whether or not our current evolutionary reconstructions and megataxonomic schemes represent an accurate outline of the organization of the virus world.

## FIGURE LEGENDS

**Figure 1.** Hierarchical taxonomic structure of the plant RNA and reverse-transcribing viromes. Only taxa containing plant-infecting viruses are shown. Genome maps of selected viruses are shown on the right. Functionally equivalent domains or genes are indicated with the same color. OuMV, Ourmia melon virus (NC_011068, NC_011069, NC_011070); CpMV1, Cryphonectria parasitica mitovirus 1 (NC_004046); BCV1, beet cryptic virus 1 (NC_011556, NC_011557); STV, southern tomato virus (NC_011591); CPMV, cowpea mosaic virus (NC_003549, NC_003550); SBMV, southern bean mosaic virus (NC_004060); TEV, tobacco etch virus (NC_001555); BdMoV, burdock mottle virus (NC_021735, NC_021736); TMV, tobacco mosaic virus (NC_001367); TYMV, turnip yellow mosaic virus (NC_004063); TBSV, tomato bushy stunt virus (NC_001554); RDV, rice dwarf virus (NC_003760, NC_003761, NC_003762, NC_003763, NC_003764, NC_003765, NC_003766, NC_003767, NC_003768, NC_003772, NC_003773, NC_003774); L-A, Saccharomyces cerevisiae virus L-A (NC_003745); TSWV, tomato spotted wilt virus (NC_002052, NC_002050, NC_002051); BYSMV, barley yellow striate mosaic cytorhabdovirus (NC_028244); CaMV, cauliflower mosaic virus (NC_001497); Athila, Arabidopsis thaliana Athila virus (X81801); Hopscotch, Zea mays Hopscotch virus (ZMU12626). Abbreviations: RdRP, RNA-dependent RNA polymerase; MP, movement protein; SJR, single jelly-roll capsid protein; i/o/f/r-CP, inner/outer/filamentous/rod-shaped capsid protein; S1/2/3H, superfamily 1/2/3 helicase; S/P-Pro, serine/cysteine protease; vOTU, viral homologue of the ovarian tumor protease; CapE, capping enzyme; En, cap-snatching endonuclease; AlkB, Alpha-ketoglutarate-dependent dioxygenase; RiS, RNA interference suppressor; GP/Gn-Gc, membrane fusion glycoprotein; M, matrix protein; P, phosphoprotein; GAG, group specific antigen; INT, integrase; RT, reverse transcriptase; RH, RNase H; LTR, long terminal repeat.

**Figure 2.** Hierarchical taxonomic structure of the phylum *Cressdnaviricota* of the ssDNA viruses. Taxa that do not contain plant viruses are indicate in grey font. Genome maps of selected viruses are shown on the right: FBNYV, Faba bean necrotic yellows virus (NC_003560, NC_003563, NC_003562, NC_003559, NC_003566, NC_003561,

NC_003564, NC_024457); MSV, maize streak virus (NC_001346); BGMV, bean golden mosaic virus (NC_004042, NC_004043). Abbreviations: Rep/RepA, rolling circle replication initiation protein; SJR, single jelly-roll capsid protein; MP, movement protein; NSP, nuclear shuttle protein.

**Figure 3.** Evolutionary scenario for the origin of plant ssDNA viruses. Geminivirids (top) and nanovirids (bottom) have evolved from two lineages of related bacterial plasmids through acquisition of the capsid protein genes (SJR1-CP and SJR2-CP, respectively) from RNA viruses on two independent occasions. The evolution of geminivirids has likely proceeded through a genomovirid-like ancestor infecting plant-pathogenic fungi or insects.

**Figure 4.** Evolutionary scenarios for the origin of two lineages of plant RNA viruses, cytorhabdovirids (top) and secovirids (bottom). Vertical evolution (co-evolution) is depicted by black arrows, whereas horizontal virus transfer from invertebrates to plants is shown with green arrows. Genome maps are not drawn to scale. Major evolutionary changes for each step are explained above the corresponding genome maps. Abbreviations: RdRP, RNA-dependent RNA polymerase; CapE, capping enzyme; Hel, helicase; Pro, protease; VP, capsid proteins; MP, movement protein; CPS/CPL, small and large capsid proteins, respectively; 32K, 32 kDa protein; G, glycoprotein; N, nucleocapsid; M, matrix protein; P, phosphoprotein. Asterisks indicate gain of additional function – RNA interference suppression – by preexisting proteins.

**Figure 5.** Hypothetical evolutionary pathways for the origin of the protist, fungal, invertebrate, vertebrate and plant RNA viromes. Dominant mechanisms of the virus lineage macroevolution including virus-host co-evolution (VHcE) and cross-species horizontal virus transmission (HVT) are depicted by arrows. Major virus taxa that emerged in each type of the organisms are listed at the left. Potential HVT pathways from protists to vertebrates and plants are not shown for the sake of clarity.

# LITERATURE CITED


1. Adkins S. 2000. Tomato spotted wilt virus-positive steps towards negative success. *Mol Plant Pathol* 1:151-7
2. Agirrezabala X, Mendez-Lopez E, Lasso G, Sanchez-Pina MA, Aranda M, Valle M. 2015. The near-atomic cryoEM structure of a flexible filamentous plant virus shows homology of its coat protein with nucleoproteins of animal viruses. *Elife* 4:e11795
3. Ahlquist P, Strauss EG, Rice CM, Strauss JH, Haseloff J, Zimmern D. 1985. Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses. *J Virol* 53:536-42
4. Atsumi G, Tomita R, Kobayashi K, Sekine KT. 2013. Prevalence and genetic diversity of an unusual virus associated with Kobu-sho disease of gentian in Japan. *J Gen Virol* 94:2360-5
5. Attoui H, Jaafar FM, Belhouchet M, de Micco P, de Lamballerie X, Brussaard CP. 2006. Micromonas pusilla reovirus: a new member of the family Reoviridae assigned to a novel proposed genus (Mimoreovirus). *J Gen Virol* 87:1375-83
6. Baulcombe D. 2004. RNA silencing in plants. *Nature* 431:356-63
7. Bekal S, Domier LL, Niblack TL, Lambert KN. 2011. Discovery and initial analysis of novel viral genomes in the soybean cyst nematode. *J Gen Virol* 92:1870-9
8. Benites LF, Poulton N, Labadie K, Sieracki ME, Grimsley N, Piganeau G. 2019. Single cell ecogenomics reveals mating types of individual cells and ssDNA viral infections in the smallest photosynthetic eukaryotes. *Philos Trans R Soc Lond B Biol Sci* 374:20190089
9. Blanc S, Gutierrez S. 2015. The specifics of vector transmission of arboviruses of vertebrates and plants. *Curr Opin Virol* 15:27-33
10. Blanc S, Uzest M, Drucker M. 2011. New research horizons in vector-transmission of plant viruses. *Curr Opin Microbiol* 14:483-91
11. Brazeau MD, Friedman M. 2015. The origin and early phylogenetic history of jawed vertebrates. *Nature* 520:490-7
12. Breitbart M, Bonnain C, Malki K, Sawaya NA. 2018. Phage puppet masters of the marine microbial realm. *Nat Microbiol* 3:754-66
13. Briddon RW, Martin DP, Roumagnac P, Navas-Castillo J, Fiallo-Olive E, et al. 2018. Alphasatellitidae: a new family with two subfamilies for the classification of geminivirus- and nanovirus-associated alphasatellites. *Arch Virol* 163:2587-600
14. Brunkard JO, Zambryski PC. 2016. Plasmodesmata enable multicellularity: new insights into their evolution, biogenesis, and functions in development and immunity. *Curr Opin Plant Biol* 35:76-83
15. Carr JP, Murphy AM, Tungadi T, Yoon JY. 2019. Plant defense signals: Players and pawns in plant-virus-vector interactions. *Plant Sci* 279:87-95
16. Chabannes M, Iskra-Caruana ML. 2013. Endogenous pararetroviruses--a reservoir of virus infection in plants. *Curr Opin Virol* 3:615-20
17. Chen Y, Dessau M, Rotenberg D, Rasmussen DA, Whitfield AE. 2019. Entry of bunyaviruses into plants and vectors. *Adv Virus Res* 104:65-96

18. Cheng S, Xian W, Fu Y, Marin B, Keller J, et al. 2019. Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. *Cell* 179:1057-67 e14
19. Chiba M, Reed JC, Prokhnevsky AI, Chapman EJ, Mawassi M, et al. 2006. Diverse suppressors of RNA silencing enhance agroinfection by a viral replicon. *Virology* 346:7-14
20. Chow CE, Suttle CA. 2015. Biogeography of Viruses in the Sea. *Annu Rev Virol* 2:41-66
21. Cobian Guemes AG, Youle M, Cantu VA, Felts B, Nulton J, Rohwer F. 2016. Viruses as Winners in the Game of Life. *Annu Rev Virol* 3:197-214
22. Coy SR, Gann ER, Pound HL, Short SM, Wilhelm SW. 2018. Viruses of Eukaryotic Algae: Diversity, Methods for Detection, and Future Directions. *Viruses* 10
23. Csorba T, Kontra L, Burgyan J. 2015. viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology* 479-480:85-103
24. Dawe AL, Nuss DL. 2013. Hypovirus molecular biology: from Koch's postulates to host self-recognition genes that restrict virus transmission. *Adv Virus Res* 86:109-47
25. Dawson WO, Bar-Joseph M, Garnsey SM, Moreno P. 2015. Citrus tristeza virus: making an ally from an enemy. *Annu Rev Phytopathol* 53:137-55
26. de Vries J, Archibald JM. 2018. Plant evolution: landmarks on the path to terrestrial life. *New Phytol* 217:1428-34
27. Delwiche CF, Cooper ED. 2015. The Evolutionary Origin of a Terrestrial Flora. *Curr Biol* 25:R899-910
28. Diemer GS, Stedman KM. 2012. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol Direct* 7:13
29. Dietzgen RG, Freitas-Astua J, Chabi-Jesus C, Ramos-Gonzalez PL, Goodin MM, et al. 2018. Dichorhaviruses in their Host Plants and Mite Vectors. *Adv Virus Res* 102:119-48
30. Dinman JD. 2012. Control of gene expression by translational recoding. *Adv Protein Chem Struct Biol* 86:129-49
31. Dolja VV. 2003. Beet yellows virus: the importance of being different. *Mol Plant Pathol* 4:91-8
32. Dolja VV, Boyko VP, Agranovsky AA, Koonin EV. 1991. Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. *Virology* 184:79-86
33. Dolja VV, Koonin EV. 2011. Common origins and host-dependent diversity of plant and animal viromes. *Curr Opin Virol* 1:322-31
34. Dolja VV, Koonin EV. 2018. Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res* 244:36-52
35. Dolja VV, Kreuze JF, Valkonen JP. 2006. Comparative and functional genomics of closteroviruses. *Virus Res* 117:38-51
36. El Baidouri M, Carpentier MC, Cooke R, Gao D, Lasserre E, et al. 2014. Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Res* 24:831-8
37. Elbeaino T, Digiaro M, Mielke-Ehret N, Muehlbach HP, Martelli GP, Ictv Report C. 2018. ICTV Virus Taxonomy Profile: Fimoviridae. *J Gen Virol* 99:1478-9
38. Falk BW, Tsai JH. 1998. Biology and molecular biology of viruses in the genus Tenuivirus. *Annu Rev Phytopathol* 36:139-63

39. Fang X, Qi Y. 2016. RNAi in Plants: An Argonaute-Centered View. *Plant Cell* 28:272-85
40. Folimonova SY, Tilsner J. 2018. Hitchhikers, highway tolls and roadworks: the interactions of plant viruses with the phloem. *Curr Opin Plant Biol* 43:82-8
41. Fuchs M, Schmitt-Keichinger C, Sanfacon H. 2017. A Renaissance in Nepovirus Research Provides New Insights Into Their Molecular Interface With Hosts and Vectors. *Adv Virus Res* 97:61-105
42. Fukuhara T. 2019. Endornaviruses: persistent dsRNA viruses with symbiotic properties in diverse eukaryotes. *Virus Genes* 55:165-73
43. Gayral P, Noa-Carrazana JC, Lescot M, Lheureux F, Lockhart BE, et al. 2008. A single Banana streak virus integration event in the banana genome as the origin of infectious endogenous pararetrovirus. *J Virol* 82:6697-710
44. Geering AD, Maumus F, Copetti D, Choisne N, Zwickl DJ, et al. 2014. Endogenous florendoviruses are major components of plant genomes and hallmarks of virus evolution. *Nat Commun* 5:5269
45. Geoghegan JL, Holmes EC. 2018. Evolutionary Virology at 40. *Genetics* 210:1151-62
46. Ghabrial SA, Caston JR, Jiang D, Nibert ML, Suzuki N. 2015. 50-plus years of fungal viruses. *Virology* 479-480:356-68
47. Gibbs AJ, Hajizadeh M, Ohshima K, Jones RAC. 2020. The Potyviruses: An Evolutionary Synthesis Is Emerging. *Viruses* 12:132
48. Gibbs AJ, Torronen M, Mackenzie AM, Wood JT, 2nd, Armstrong JS, et al. 2011. The enigmatic genome of Chara australis virus. *J Gen Virol* 92:2679-90
49. Gilmer D, Ratti C, Ictv Report C. 2017. ICTV Virus Taxonomy Profile: Benyviridae. *J Gen Virol* 98:1571-2
50. Goldbach R, Wellink J. 1988. Evolution of plus-strand RNA viruses. *Intervirology* 29:260-7
51. Granot D, Kelly G. 2019. Evolution of Guard-Cell Theories: The Story of Sugars. *Trends Plant Sci* 24:507-18
52. Gray S, Gildow FE. 2003. Luteovirus-aphid interactions. *Annu Rev Phytopathol* 41:539-66
53. Gregory AC, Zayed AA, Conceicao-Neto N, Temperton B, Bolduc B, et al. 2019. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell* 177:1109-23 e14
54. Greninger AL. 2018. A decade of RNA virus metagenomics is (not) enough. *Virus Res* 244:218-29
55. Grigoras I, Vetten HJ, Commandeur U, Ziebell H, Gronenborn B, Timchenko T. 2018. Nanovirus DNA-N encodes a protein mandatory for aphid transmission. *Virology* 522:281-91
56. Gronenborn B. 2004. Nanoviruses: genome organisation and protein function. *Vet Microbiol* 98:103-9
57. Grybchuk D, Akopyants NS, Kostygov AY, Konovalovas A, Lye LF, et al. 2018. Viral discovery and diversity in trypanosomatid protozoa with a focus on relatives of the human parasite Leishmania. *Proc Natl Acad Sci U S A* 115:E506-E15
58. Grybchuk D, Kostygov AY, Macedo DH, d'Avila-Levy CM, Yurchenko V. 2018. RNA viruses in trypanosomatid parasites: a historical overview. *Mem Inst Oswaldo Cruz* 113:e170487
59. Guiry MD. 2012. How Many Species of Algae Are There? *J Phycol* 48:1057-63
60. Guo Z, Li Y, Ding SW. 2019. Small RNA-based antimicrobial immunity. *Nat Rev Immunol* 19:31-44

61. Hamelin FM, Allen LJ, Prendeville HR, Hajimorad MR, Jeger MJ. 2016. The evolution of plant virus transmission pathways. *J Theor Biol* 396:75-89
62. Heinlein M. 2015. Plant virus replication and movement. *Virology* 479-480:657-71
63. Hesketh EL, Saunders K, Fisher C, Potze J, Stanley J, et al. 2018. The 3.3 A structure of a plant geminivirus using cryo-EM. *Nat Commun* 9:2369
64. Hillman BI, Cai G. 2013. The family narnaviridae: simplest of RNA viruses. *Adv Virus Res* 86:149-76
65. Hogenhout SA, Ammar el D, Whitfield AE, Redinbaugh MG. 2008. Insect vector interactions with persistently transmitted viruses. *Annu Rev Phytopathol* 46:327-59
66. Hohn T, Rothnie H. 2013. Plant pararetroviruses: replication and expression. *Curr Opin Virol* 3:621-8
67. Iranzo J, Koonin EV, Prangishvili D, Krupovic M. 2016. Bipartite Network Analysis of the Archaeal Virosphere: Evolutionary Connections between Viruses and Capsidless Mobile Elements. *J Virol* 90:11043-55
68. Iranzo J, Krupovic M, Koonin EV. 2016. The Double-Stranded DNA Virosphere as a Modular Hierarchical Network of Gene Sharing. *MBio* 7
69. Jiang D, Ayllon MA, Marzano SL, Ictv Report C. 2019. ICTV Virus Taxonomy Profile: Mymonaviridae. *J Gen Virol* 100:1343-4
70. Jones JD, Vance RE, Dangl JL. 2016. Intracellular innate immune surveillance devices in plants and animals. *Science* 354
71. Kamer G, Argos P. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res* 12:7269-82
72. Kanyuka K, Ward E, Adams MJ. 2003. Polymyxa graminis and the cereal viruses it transmits: a research challenge. *Mol Plant Pathol* 4:393-406
73. Karasev AV. 2000. Genetic Diversity and Evolution of Closteroviruses. *Annu Rev Phytopathol* 38:293-324
74. Kazlauskas D, Varsani A, Koonin EV, Krupovic M. 2019. Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nat Commun* 10:3425
75. Kazlauskas D, Varsani A, Krupovic M. 2018. Pervasive Chimerism in the Replication-Associated Proteins of Uncultured Single-Stranded DNA Viruses. *Viruses* 10
76. Keeling PJ, Burki F. 2019. Progress towards the Tree of Eukaryotes. *Curr Biol* 29:R808-R17
77. Kenrick P, Wellman CH, Schneider H, Edgecombe GD. 2012. A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philos Trans R Soc Lond B Biol Sci* 367:519-36
78. Kirsip H, Abroi A. 2019. Protein Structure-Guided Hidden Markov Models (HMMs) as A Powerful Method in the Detection of Ancestral Endogenous Viral Elements. *Viruses* 11
79. Kobayashi K, Atsumi G, Iwadate Y, Tomita R, Chiba K, et al. 2013. Gentian Kobu-sho-associated virus: a tentative, novel double-stranded RNA virus that is relevant to gentian Kobu-sho syndrome. *J. Gen. Plant Pathol.* 79:56-63
80. Koga R, Fukuhara T, Nitta T. 1998. Molecular characterization of a single mitochondria-associated double-stranded RNA in the green alga Bryopsis. *Plant Mol Biol* 36:717-24

81. Koga R, Horiuchi H, Fukuhara T. 2003. Double-stranded RNA replicons associated with chloroplasts of a green alga, Bryopsis cinicola. *Plant Mol Biol* 51:991-9
82. Koonin EV, Dolja VV. 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol* 28:375-430
83. Koonin EV, Dolja VV. 2014. Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol Mol Biol Rev* 78:278-303
84. Koonin EV, Dolja VV, Krupovic M. 2015. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* 479-480:2-25
85. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, et al. 2020. Global Organization and Proposed Megataxonomy of the Virus World. *Micobiol. Mol. Biol. Rev.*:in press
86. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, et al. 2020. Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.*:in press
87. Koonin EV, Krupovic M. 2018. The depths of virus exaptation. *Curr Opin Virol* 31:1-8
88. Koonin EV, Wolf YI, Katsnelson MI. 2017. Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states. *Biol Direct* 12:31
89. Koonin EV, Wolf YI, Nagasaki K, Dolja VV. 2008. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nat Rev Microbiol* 6:925-39
90. Kormelink R, Garcia ML, Goodin M, Sasaya T, Haenni AL. 2011. Negative-strand RNA viruses: the plant-infecting counterparts. *Virus Res* 162:184-202
91. Kotta-Loizou I, Coutts RH. 2017. Studies on the Virome of the Entomopathogenic Fungus Beauveria bassiana Reveal Novel dsRNA Elements and Mild Hypervirulence. *PLoS Pathog* 13:e1006183
92. Krupovic M. 2013. Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. *Curr Opin Virol* 3:578-86
93. Krupovic M, Blomberg J, Coffin JM, Dasgupta I, Fan H, et al. 2018. Ortervirales: A new viral order unifying five families of reverse-transcribing viruses. *J Virol*
94. Krupovic M, Dolja VV, Koonin EV. 2015. Plant viruses of the Amalgaviridae family evolved via recombination between viruses with double-stranded and negative-strand RNA genomes. *Biol Direct* 10:12
95. Krupovic M, Dolja VV, Koonin EV. 2019. Origin of viruses: primordial replicators recruiting capsids from hosts. *Nat Rev Microbiol* 17:449-58
96. Krupovic M, Ghabrial SA, Jiang D, Varsani A. 2016. Genomoviridae: a new family of widespread single-stranded DNA viruses. *Arch Virol* 161:2633-43
97. Krupovic M, Koonin EV. 2017. Homologous Capsid Proteins Testify to the Common Ancestry of Retroviruses, Caulimoviruses, Pseudoviruses, and Metaviruses. *J Virol* 91
98. Krupovic M, Koonin EV. 2017. Multiple origins of viral capsid proteins from cellular ancestors. *Proc Natl Acad Sci U S A* 114:E2401-E10
99. Krupovic M, Ravantti JJ, Bamford DH. 2009. Geminiviruses: a tale of a plasmid becoming a virus. *BMC Evol Biol* 9:112
100. Lachnit T, Thomas T, Steinberg P. 2016. Expanding our Understanding of the Seaweed Holobiont: RNA Viruses of the Red Alga Delisea pulchra. *Front Microbiol* 6:1489

101. Lakatos L, Csorba T, Pantaleo V, Chapman EJ, Carrington JC, et al. 2006. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J* 25:2768-80
102. Laten HM, Majumdar A, Gaucher EA. 1998. SIRE-1, a copia/Ty1-like retroelement from soybean, encodes a retroviral envelope-like protein. *Proc Natl Acad Sci U S A* 95:6897-902
103. Leastro MO, Kitajima EW, Silva MS, Resende RO, Freitas-Astua J. 2018. Dissecting the Subcellular Localization, Intracellular Trafficking, Interactions, Membrane Association, and Topology of Citrus Leprosis Virus C Proteins. *Front Plant Sci* 9:1299
104. Lecoq H, Wipf-Scheibel C, Verdin E, Desbiez C. 2019. Characterization of the first tenuivirus naturally infecting dicotyledonous plants. *Arch Virol* 164:297-301
105. Lefeuvre P, Martin DP, Elena SF, Shepherd DN, Roumagnac P, Varsani A. 2019. Evolution and ecology of plant viruses. *Nat Rev Microbiol*
106. Li CX, Shi M, Tian JH, Lin XD, Kang YJ, et al. 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* 4
107. Li F, Wang A. 2019. RNA-Targeted Antiviral Immunity: More Than Just RNA Silencing. *Trends Microbiol* 27:792-805
108. Lin YH, Fujita M, Chiba S, Hyodo K, Andika IB, et al. 2019. Two novel fungal negative-strand RNA viruses related to mymonaviruses and phenuiviruses in the shiitake mushroom (Lentinula edodes). *Virology* 533:125-36
109. Liu H, Fu Y, Jiang D, Li G, Xie J, et al. 2009. A novel mycovirus that is related to the human pathogen hepatitis E virus and rubi-like viruses. *J Virol* 83:1981-91
110. Liu L, Chen X. 2018. Intercellular and systemic trafficking of RNAs in plants. *Nat Plants* 4:869-78
111. Llorens C, Futami R, Covelli L, Dominguez-Escriba L, Viu JM, et al. 2011. The Gypsy Database (GyDB) of mobile genetic elements: release 2.0. *Nucleic Acids Res* 39:D70-4
112. Llorens C, Munoz-Pomer A, Bernad L, Botella H, Moya A. 2009. Network dynamics of eukaryotic LTR retroelements beyond phylogenetic trees. *Biol Direct* 4:41
113. Lu G, Li S, Zhou C, Qian X, Xiang Q, et al. 2019. Tenuivirus utilizes its glycoprotein as a helper component to overcome insect midgut barriers for its circulative and propagative transmission. *PLoS Pathog* 15:e1007655
114. Lubicz JV, Rush CM, Payton M, Colberg T. 2007. Beet necrotic yellow vein virus accumulates inside resting spores and zoosporangia of its vector Polymyxa betae BNYVV infects P. betae. *Virol J* 4:37
115. Malik HS, Henikoff S, Eickbush TH. 2000. Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* 10:1307-18
116. Martelli GP, Adams MJ, Kreuze JF, Dolja VV. 2007. Family Flexiviridae: a case study in virion and genome plasticity. *Annu Rev Phytopathol* 45:73-100
117. Marzano SY, Nelson BD, Ajayi-Oyetunde O, Bradley CA, Hughes TJ, et al. 2016. Identification of Diverse Mycoviruses through Metatranscriptomics Characterization of the Viromes of Five Major Fungal Plant Pathogens. *J Virol* 90:6846-63
118. Mata CP, Luque D, Gomez-Blanco J, Rodriguez JM, Gonzalez JM, et al. 2017. Acquisition of functions on the outer capsid surface during evolution of double-stranded RNA fungal viruses. *PLoS Pathog* 13:e1006755

119. Maumus F, Epert A, Nogue F, Blanc G. 2014. Plant genomes enclose footprints of past infections by giant virus relatives. *Nat Commun* 5:4268
120. McLeish MJ, Fraile A, Garcia-Arenal F. 2019. Evolution of plant-virus interactions: host range and virus emergence. *Curr Opin Virol* 34:50-5
121. Melnyk CW, Molnar A, Baulcombe DC. 2011. Intercellular and systemic movement of RNA silencing signals. *EMBO J* 30:3553-63
122. Misof B, Liu S, Meusemann K, Peters RS, Donath A, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763-7
123. Miyazaki N, Nakagawa A, Iwasaki K. 2013. Life cycle of phytoreoviruses visualized by electron microscopy and tomography. *Front Microbiol* 4:306
124. Mushegian A, Shipunov A, Elena SF. 2016. Changes in the composition of the RNA virome mark evolutionary transitions in green plants. *BMC Biol* 14:68
125. Mushegian AR, Elena SF. 2015. Evolution of plant virus movement proteins from the 30K superfamily and of their homologs integrated in plant genomes. *Virology* 476:304-15
126. Mushegian AR, Koonin EV. 1993. Cell-to-cell movement of plant viruses. Insights from amino acid sequence comparisons of movement proteins and from analogies with cellular transport systems. *Arch Virol* 133:239-57
127. Napuli AJ, Alzhanova DV, Doneanu CE, Barofsky DF, Koonin EV, Dolja VV. 2003. The 64-kilodalton capsid protein homolog of Beet yellows virus is required for assembly of virion tails. *J Virol* 77:2377-84
128. Naranjo-Ortiz MA, Gabaldon T. 2019. Fungal evolution: major ecological adaptations and evolutionary transitions. *Biol Rev Camb Philos Soc* 94:1443-76
129. Navarro JA, Sanchez-Navarro JA, Pallas V. 2019. Key checkpoints in the movement of plant viruses through the host. *Adv Virus Res* 104:1-64
130. Nawaz-ul-Rehman MS, Fauquet CM. 2009. Evolution of geminiviruses and their satellites. *FEBS Lett* 583:1825-32
131. Nerva L, Forgia M, Ciuffo M, Chitarra W, Chiapello M, et al. 2019. The mycovirome of a fungal collection from the sea cucumber Holothuria polii. *Virus Res* 273:197737
132. Nerva L, Turina M, Zanzotto A, Gardiman M, Gaiotti F, et al. 2019. Isolation, molecular characterization and virome analysis of culturable wood fungal endophytes in esca symptomatic and asymptomatic grapevine plants. *Environ Microbiol* 21:2886-904
133. Neuhauser S, Kirchmair M, Bulman S, Bass D. 2014. Cross-kingdom host shifts of phytomyxid parasites. *BMC Evol Biol* 14:33
134. Neumann P, Novak P, Hostakova N, Macas J. 2019. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. *Mob DNA* 10:1
135. Ng JC, Falk BW. 2006. Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annu Rev Phytopathol* 44:183-212
136. Nibert ML, Debat HJ, Manny AR, Grigoriev IV, De Fine Licht HH. 2019. Mitovirus and Mitochondrial Coding Sequences from Basal Fungus Entomophthora muscae. *Viruses* 11
137. Nibert ML, Ghabrial SA, Maiss E, Lesker T, Vainio EJ, et al. 2014. Taxonomic reorganization of family Partitiviridae and other recent progress in partitivirus research. *Virus Res* 188:128-41

138. Nibert ML, Vong M, Fugate KK, Debat HJ. 2018. Evidence for contemporary plant mitoviruses. *Virology* 518:14-24
139. Niklas KJ, Cobb ED, Matas AJ. 2017. The evolution of hydrophobic cell wall biopolymers: from algae to angiosperms. *J Exp Bot* 68:5261-9
140. Niklas KJ, Newman SA. 2019. The Many Roads To (and From) Multicellularity. *J Exp Bot*
141. Novikova O, Belfort M. 2017. Mobile Group II Introns as Ancestral Eukaryotic Elements. *Trends Genet* 33:773-83
142. Novikova O, Smyshlyaev G, Blinov A. 2010. Evolutionary genomics revealed interkingdom distribution of Tcn1-like chromodomain-containing Gypsy LTR retrotransposons among fungi and plants. *BMC Genomics* 11:231
143. One Thousand Plant Transcriptomes I. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574:679-85
144. Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, et al. 2016. Uncovering Earth's virome. *Nature* 536:425-30
145. Pascon RC, Kitajima JP, Breton MC, Assumpcao L, Greggio C, et al. 2006. The complete nucleotide sequence and genomic organization of Citrus Leprosis associated Virus, Cytoplasmatic type (CiLV-C). *Virus Genes* 32:289-98
146. Peterson KJ, Cotton JA, Gehling JG, Pisani D. 2008. The Ediacaran emergence of bilaterians: congruence between the genetic and the geological fossil records. *Philos Trans R Soc Lond B Biol Sci* 363:1435-43
147. Popper ZA, Michel G, Herve C, Domozych DS, Willats WG, et al. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu Rev Plant Biol* 62:567-90
148. Pyle JD, Keeling PJ, Nibert ML. 2017. Amalga-like virus infecting Antonospora locustae, a microsporidian pathogen of grasshoppers, plus related viruses associated with other arthropods. *Virus Res* 233:95-104
149. Quito-Avila DF, Brannen PM, Cline WO, Harmon PF, Martin RR. 2013. Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. *J Gen Virol* 94:1426-34
150. Rastgou M, Habibi MK, Izadpanah K, Masenga V, Milne RG, et al. 2009. Molecular characterization of the plant virus genus Ourmiavirus and evidence of inter-kingdom reassortment of viral genome segments as its possible route of origin. *J Gen Virol* 90:2525-35
151. Revers F, Garcia JA. 2015. Molecular biology of potyviruses. *Adv Virus Res* 92:101-99
152. Richert-Poggeler KR, Noreen F, Schwarzacher T, Harper G, Hohn T. 2003. Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. *EMBO J* 22:4836-45
153. Richert-Poggeler KR, Shepherd RJ. 1997. Petunia vein-clearing virus: a plant pararetrovirus with the core sequences for an integrase function. *Virology* 236:137-46
154. Rochon D, Kakani K, Robbins M, Reade R. 2004. Molecular aspects of plant virus transmission by olpidium and plasmodiophorid vectors. *Annu Rev Phytopathol* 42:211-41
155. Roossinck MJ. 2010. Lifestyles of plant viruses. *Philos Trans R Soc Lond B Biol Sci* 365:1899-905

156. Roossinck MJ. 2012. Plant virus metagenomics: biodiversity and ecology. *Annu Rev Genet* 46:359-69
157. Roossinck MJ. 2018. Evolutionary and ecological links between plant and fungal viruses. *New Phytol*
158. Roossinck MJ, Sabanadzovic S, Okada R, Valverde RA. 2011. The remarkable evolutionary history of endornaviruses. *J Gen Virol* 92:2674-8
159. Rosa C, Kuo YW, Wuriyanghan H, Falk BW. 2018. RNA Interference Mechanisms and Applications in Plant Pathology. *Annu Rev Phytopathol* 56:581-610
160. Rota-Stabelli O, Daley AC, Pisani D. 2013. Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Curr Biol* 23:392-8
161. Rousvoal S, Bouyer B, Lopez-Cristoffanini C, Boyen C, Collen J. 2016. Mutant swarms of a totivirus-like entities are present in the red macroalga Chondrus crispus and have been partially transferred to the nuclear genome. *J Phycol* 52:493-504
162. Roux S, Enault F, Bronner G, Vaulot D, Forterre P, Krupovic M. 2013. Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses. *Nat Commun* 4:2700
163. Sabanadzovic S, Valverde RA, Brown JK, Martin RR, Tzanetakis IE. 2009. Southern tomato virus: The link between the families Totiviridae and Partitiviridae. *Virus Res* 140:130-7
164. Sanfacon H, Wellink J, Le Gall O, Karasev A, van der Vlugt R, Wetzel T. 2009. Secoviridae: a proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae, the unassigned genera Cheravirus and Sadwavirus, and the proposed genus Torradovirus. *Arch Virol* 154:899-907
165. Scholthof KB. 2004. Tobacco mosaic virus: a model system for plant biology. *Annu Rev Phytopathol* 42:13-34
166. Scholthof KB, Adkins S, Czosnek H, Palukaitis P, Jacquot E, et al. 2011. Top 10 plant viruses in molecular plant pathology. *Mol Plant Pathol* 12:938-54
167. Shi M, Lin XD, Chen X, Tian JH, Chen LJ, et al. 2018. The evolutionary history of vertebrate RNA viruses. *Nature* 556:197-202
168. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, et al. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539-43
169. Shi M, Lin XD, Vasilakis N, Tian JH, Li CX, et al. 2016. Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. *J Virol* 90:659-69
170. Sicard A, Michalakis Y, Gutierrez S, Blanc S. 2016. The Strange Lifestyle of Multipartite Viruses. *PLoS Pathog* 12:e1005819
171. Sicard A, Zeddam JL, Yvon M, Michalakis Y, Gutierrez S, Blanc S. 2015. Circulative Nonpropagative Aphid Transmission of Nanoviruses: an Oversimplified View. *J Virol* 89:9719-26
172. Sobhy H. 2017. A comparative review of viral entry and attachment during large and giant dsDNA virus infections. *Arch Virol* 162:3567-85
173. Soltis PS, Folk RA, Soltis DE. 2019. Darwin review: angiosperm phylogeny and evolutionary radiations. . *Proc. R. Soc. B* 286:20190099

174. Somera M, Sarmiento C, Truve E. 2015. Overview on Sobemoviruses and a Proposal for the Creation of the Family Sobemoviridae. *Viruses* 7:3076-115
175. Spatafora JW, Aime MC, Grigoriev IV, Martin F, Stajich JE, Blackwell M. 2017. The Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies. *Microbiol Spectr* 5
176. Stork NE. 2018. How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth? *Annu Rev Entomol* 63:31-45
177. Taliansky M, Mayo MA, Barker H. 2003. Potato leafroll virus: a classic pathogen shows some new tricks. *Mol Plant Pathol* 4:81-9
178. Teixeira M, Sela N, Ng J, Casteel CL, Peng HC, et al. 2016. A novel virus from Macrosiphum euphorbiae with similarities to members of the family Flaviviridae. *J Gen Virol* 97:1261-71
179. Toriyama S, Kimishima T, Takahashi M, Shimizu T, Minaka N, Akutsu K. 1998. The complete nucleotide sequence of the rice grassy stunt virus genome and genomic comparisons with viruses of the genus Tenuivirus. *J Gen Virol* 79 ( Pt 8):2051-8
180. Van Etten JL, Agarkova IV, Dunigan DD. 2019. Chloroviruses. *Viruses* 12
181. Vasilakis N, Forrester NL, Palacios G, Nasar F, Savji N, et al. 2013. Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. *J Virol* 87:2475-88
182. Verchot-Lubicz J, Torrance L, Solovyev AG, Morozov SY, Jackson AO, Gilmer D. 2010. Varied movement strategies employed by triple gene block-encoding viruses. *Mol Plant Microbe Interact* 23:1231-47
183. Vlok M, Gibbs AJ, Suttle CA. 2019. Metagenomes of a Freshwater Charavirus from British Columbia Provide a Window into Ancient Lineages of Viruses. *Viruses* 11
184. Wei T, Li Y. 2016. Rice Reoviruses in Insect Vectors. *Annu Rev Phytopathol* 54:99-120
185. Weynberg KD, Allen MJ, Wilson WH. 2017. Marine Prasinoviruses and Their Tiny Plankton Hosts: A Review. *Viruses* 9
186. White KA, Nagy PD. 2004. Advances in the molecular biology of tombusviruses: gene expression, genome replication, and recombination. *Prog Nucleic Acid Res Mol Biol* 78:187-226
187. Whitfield AE, Falk BW, Rotenberg D. 2015. Insect vector-mediated transmission of plant viruses. *Virology* 479-480:278-89
188. Whitfield AE, Huot OB, Martin KM, Kondo H, Dietzgen RG. 2018. Plant rhabdoviruses-their origins and vector interactions. *Curr Opin Virol* 33:198-207
189. Wolf YI, Kazlauskas D, Iranzo J, Lucia-Sanz A, Kuhn JH, et al. 2018. Origins and evolution of the global RNA virome. *mBio* 9:e02329-18
190. Wood R, Liu AG, Bowyer F, Wilby PR, Dunn FS, et al. 2019. Integrated records of environmental change and evolution challenge the Cambrian Explosion. *Nat Ecol Evol* 3:528-38
191. Wright DA, Voytas DF. 1998. Potential retroviruses in plants: Tat1 is related to a group of Arabidopsis thaliana Ty3/gypsy retrotransposons that encode envelope-like proteins. *Genetics* 149:703-15
192. Yu X, Li B, Fu Y, Xie J, Cheng J, et al. 2013. Extracellular transmission of a DNA mycovirus and its use as a natural fungicide. *Proc Natl Acad Sci U S A* 110:1452-7

193. Zamora M, Mendez-Lopez E, Agirrezabala X, Cuesta R, Lavin JL, et al. 2017. Potyvirus virion structure shows conserved protein fold and RNA binding site in ssRNA viruses. *Sci Adv* 3:eaao2182
194. Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, et al. 2017. ICTV Virus Taxonomy Profile: Geminiviridae. *J Gen Virol* 98:131-3
195. Zhang YZ, Chen YM, Wang W, Qin XC, Holmes EC. 2019. Expanding the RNA Virosphere by Unbiased Metagenomics. *Annu Rev Virol*
196. Zhang YZ, Wu WC, Shi M, Holmes EC. 2018. The diversity, evolution and origins of vertebrate RNA viruses. *Curr Opin Virol* 31:9-16
197. Zhao L, Rosario K, Breitbart M, Duffy S. 2019. Eukaryotic Circular Rep-Encoding Single-Stranded DNA (CRESS DNA) Viruses: Ubiquitous Viruses With Small Genomes and a Diverse Host Range. *Adv Virus Res* 103:71-133