

A novel weevil-transmitted tymovirus found in mixed infection on hollyhock

Mathieu Mahillon

Research Group in Virology, Bacteriology and Phytoplasmaology, Agroscope

Justine Brodard

Research Group in Virology, Bacteriology and Phytoplasmaology, Agroscope

Isabelle Kellenberger

Research Group in Virology, Bacteriology and Phytoplasmaology, Agroscope

Arnaud Blouin

Research Group in Virology, Bacteriology and Phytoplasmaology, Agroscope

Olivier Schumpp (✉ olivier.schumpp@agroscope.admin.ch)

Research Group in Virology, Bacteriology and Phytoplasmaology, Agroscope

Research Article

Keywords: tymovirus, tymoviridae, weevil, hollyhock, mixed infection, carlavirus, potyvirus

Posted Date: November 23rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-2285443/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Virology Journal on January 30th, 2023.
See the published version at <https://doi.org/10.1186/s12985-023-01976-6>.

Abstract

Leaves of hollyhock (*Alcea rosea*) exhibiting vein chlorosis and yellow mosaic symptoms were collected at public sites in Lausanne and Nyon, two cities of western Switzerland. Classical diagnostic methods untangled in samples from both sites the mixed infections of a novel isometric virus, tentatively named “Alcea yellow mosaic virus” (AYMV) with the carlavirus Gaillardia latent virus. A new potyvirus was also identified in samples from Nyon. A combination of Illumina, Nanopore and Sanger sequencing was necessary to assemble the full-length genome of AYMV, revealing an exceptionally high cytidine content and other features typically associated with members of the genus *Tymovirus*. The host range of AYMV was found to be restricted to mallows, including ornamentals as well as economically important plants. Phylogenetic analyses further showed that AYMV belongs to a *Tymovirus* subclade that also gathers the other mallow-infecting members. The virus was readily transmitted by sap inoculation, and the weevil species *Aspidapion radiolus* was evidenced as a vector. Transmission assays using another weevil or other insect species did not succeed, and seed transmission was not observed.

Introduction

The genus *Tymovirus* (*Tymovirales*: *Tymoviridae*) accommodates isometric viruses infecting dicotyledons (1). Historically, members of this taxon have served as models for studies of virion structure, transfer RNA (tRNA) mimicry and RNA silencing (2–5). Several members are also pathogens of economically important crops. For instance, the type member turnip yellow mosaic virus (TYMV) infects diverse *Brassica* plants in Eurasia and Australia (6, 7). Another example is tomato blistering mosaic virus (TBMV), an emerging threat to tomato and tobacco production in South America (8, 9). In parallel, some members have been identified on non-crop plants, revealing interesting aspects of viral ecology in wild natural habitats (10–12). Tymoviruses are often found at high titer in plant tissues and typically induce yellow mosaic symptoms on leaves. Transmission can be mediated via chewing insects from the order *Coleoptera* (1), plant-to-plant contact, agricultural machineries, and in rare cases through seeds (13, 14).

The genomic RNA (gRNA) of tymoviruses is a positive-sense, single-stranded RNA (+ ssRNA) that ranges from 6,035 to 6,679 nt in length (1). It harbors a 5' cap and ends by a tRNA-like structure (TLS) with a few exceptions (15, 16). Most molecular data regarding the tymoviruses have been collected from TYMV for which the model species *Arabidopsis thaliana* is a suitable host (17). Three ORFs are present on the TYMV genome; two overlapping ORFs encoding for the movement protein (MP) and the replication protein (RP) followed by small ORF encoding for the capsid protein (CP) (1). The MP acts as a viral suppressor of RNA silencing (VSR) (18) and is involved in symptom induction (19). The RP, translated by leaky scanning, is a polyprotein displaying activities of methyltransferase (Met), papain-like protease (Pro), helicase (Hel) and RNA-dependent RNA polymerase (RdRp). Following two Pro-mediated cleavages, the RP produces three mature peptides that enable replication in chloroplast-associated vesicles (20, 21). The CP ORF is translated from a subgenomic RNA (sgRNA) for which the promoter is a conserved sequence named the “tymobox” (22).

The family *Malvaceae*, commonly referred to as the mallows, is a large group of flowering plants which comprises several species of economic importance. To date, two tymoviruses have been retrieved from mallows, namely okra mosaic virus (OMV) and cacao yellow mosaic virus (CYMV), both present in Sub-Saharan Africa (23–25). The mallows also include numerous ornamentals such as the hollyhock, a popular biennial plant commonly grown in private gardens and public areas. In this study, we present a novel tymovirus retrieved from symptomatic hollyhocks in two cities of western Switzerland.

Material And Method

Symptomatic plant material

Leaf samples from symptomatic hollyhocks were collected in two green parks in Lausanne (46°31'10.9"N 6°35'57.1"E) and Nyon (46°23'05.4"N 6°14'06.4"E) in May 2021 and were stored at 4°C until further use.

Mechanical inoculation

The leaf material was ground in cold phosphate buffer (Na_2HPO_4 20 mM, pH 7.6, supplemented with 1 mM diethyldithiocarbamate) in a mortar using a pestle. The sap was then rubbed on the first true leaves of tested plants using carborundum (400-mesh silicon carbide, Aldrich) as abrasive. Plants were inoculated when the first leaves were fully developed and then maintained for at least two months in greenhouse conditions (24°C, 14/10h photoperiod) with daily inspection for symptom expression. Inoculation of semi-purified particles was performed in a similar way.

Semi-purification of virions

A protocol developed for nepoviral particles was followed with some modifications (26). Briefly, leaves were first ground in liquid N_2 using a mortar and a pestle. The obtained powder was mixed with a cold extraction buffer (17 mM citric acid, 167 mM NaH_2PO_4 , 0.2% thioglycolic acid, pH 7.0) in a ratio 1:4 (w/v). The solution was slowly mixed on ice for 20 min and then filtered twice through cheesecloth. The filtrate was next clarified with 0.5 vol chloroform and 0.5 vol butanol. After slow mixing on ice for 20 min, the solution was centrifuged at 4,000 rpm for 20 min. The supernatant was collected and centrifuged for 1h at 40 000 rpm. The obtained pellet was resuspended in 1:10 diluted extraction buffer. Semi-purified particles were denatured in Laemmli buffer and then analysed by SDS-PAGE using a precast 12% acrylamide gel (Bio-rad). An unstained, SDS-PAGE standards low range ladder (Biorads) was used for size estimation. The gel was stained with coomassie brilliant blue R 250 (Fluka). Total RNA was extracted from semi-purified particles via the RNeasy plant kit (Qiagen, with DNase I treatment), and submitted to an agarose gel electrophoresis. For this gel as well as other agarose gels, a 100bp DNA ladder was used.

Transmission electron microscopy

Leaf crude extracts and semi-purified particles were analyzed by transmission electron microscopy (TEM). Samples were first mixed with one vol of 0.1% bovine serum albumin and one vol of 4% phosphotungstic acid (pH 6.0). One drop of the resulting solution was sprayed onto Formvar/Carbon 400-

Mesh copper grids (Agar scientific) using a homemade device. Grids were then observed with the Tecnai G2 Spirit electron microscope (FEI).

RT-PCR and Sanger sequencing

RT-PCR analyses were performed on CTAB-extracted nucleic acids as previously described (27) using the primers listed in Table 1. When necessary, amplicons were cloned in the pGemT plasmid using T4 DNA ligase (Promega). Recombinant plasmids were then used to transform *E. coli* DH5alpha and subsequently sequenced by Fasteris (Switzerland) or Microsynth (Switzerland) using M13 universal primers.

Table 1
Primers used in this study.

Name	Sequence	Use	Reference
CIFor	GGIVVIGTIGGIWSIGGIAARTCIAC	Potyvirus generic detection	(28)
CIRev	ACICCRTTYTCDATDATRTTIGTIGC		
TymZ-F	GGSCCMGTSAARAARTAYCA	Tymovirus generic detection	(29)
TymZ-R	GCCAGRTTGTARTCRGRGTTG		
CarF2b	GGRCTDGGDGTVCCNACTGA	Carlavirus generic detection	(30)
CarR	CCWCATYSRCTCSRCTWTGG		
TymoFinalF	CCCTACACCACGTCCCATTTC	AYMV assembly gap & detection	This study
TymoFinalR	CAAGTGGGGTCGGTCAGAAG		
Seq5two	GATGGCCCAGGGAAGTTCTG	AYMV 5' terminus	
Tymo5closer	TCCGTGACCATCTCTGGAGT	AYMV 3' terminus	
AnchoredPolyA	A ₂₂ BNN		
Carla_GLV_For	GGTACAGGAATTGGCTGGCT	GLV detection	
Carla_GLV_Rev	TTCCAATACTGACGGCTCGG		
Poty_Malva_Fwd	GAACCAACACGACCACTTGC	Novel potyvirus detection	
Poty_Malva_Rv	AGTGTGGGTTGGCTGAAGTT		
I = inosine, Y = C/T, B = C/G/T, R = G/A, W = A/T, V = A/C/G, S = C/G and D = A/G/T			

Extraction of dsRNA

A cellulose-based protocol was followed to extract the viral double-stranded RNA (dsRNA) from infected leaves (31). The dsRNA was resuspended in water and then treated with DNase QI (Qiagen) and S1 nuclease (Promega) to remove plant DNA and RNA.

Genome sequencing

The dsRNA extracted from infected *M. sylvestris* plants was first quantified using the Qubit 3 Fluorometer (Invitrogen), and then sequenced by high-throughput sequencing (HTS) on a Illumina platform at MacroGen (South Korea), using a paired-end, 100 bp-read cDNA library. A total of 32,314,372 reads were generated, of which > 95.2% exhibited quality scores > Q30. Reads were trimmed with Trimmomatic (32) and assembled using Spades (33). Viral contigs were then sorted based on Blastx and Blastn searches in the NCBI databases limited to the taxid “virus”. Gaps between viral contigs were filled by RT-PCR and subsequent Sanger sequencing, except for one region that was sequenced on a MinION Flongle flow cell (Oxford Nanopore technologies) using the Ligation Sequencing Kit SQK-LSK110. The flow cell ran with 75 active sites and produced a total of 216,826 reads. Short and long reads were mapped on the viral contigs using Bowtie2 (34) and Minimap2 (35), respectively, and coverages were determined using SAMtools (36).

The viral termini were first amplified by RT-PCR on CTAB-extracted nucleic acids and then obtained by Sanger sequencing. For the 5′ terminus, the SMARTer RACE 5′/3′ Kit (Takara) was used in combination with the primer Seq5Two (Table 1). For the 3′ terminus, a poly(U) tailing was first performed using the Poly(U) Polymerase (NEB). The viral 3′ end was then amplified by RT-PCR as previously described (26) using the primers Tymo5closer and AnchoredPolyA (Table 1). For both termini, amplicons were cloned and sequenced as mentioned above.

Bioinformatics analyses

Viral sequences were analyzed on Ugene (37) and Jalview (38). Conserved protein domains were identified with the CDD database (39). RNA secondary structures were predicted as previously described (40). Proteins and nucleotides were aligned with Muscle (41). Pairwise identities matrix for these alignments were generated with the SDT software (42). Amino acid (AA) contents were obtained with EMBOSS pepstats (available at: https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/). AA logos were produced on Seq2logo (43). For phylogenetic analyses, the best fit substitution models for the protein alignments were first retrieved with ModelFinder (44) and used to create maximum likelihood (ML) phylogenetic trees on IQ-Tree (45) in combination with ultrafast bootstraps (46). ML trees were curated on ITol (47).

Transmission assays

For the insect transmission assays, specimens were placed for 24 h in homemade plastic boxes containing detached symptomatic leaves of *M. sylvestris*. Then, pairs of insects were placed onto 3-week old *M. sylvestris* plants covered in insect-proof nets (Andermatt), and carefully removed after one week. For transmission assays using aphids, plants were also treated with the commercial insecticide Tapeki. Plants were then maintained one month in greenhouse for symptom observation. For each plant, the infection status was then determined by RT-PCR.

From one infected plant in Lausanne, seeds were collected in the 2021 autumn and kept at 4°C until the following spring. Seeds were then sown and the resulting plants were grown in greenhouse conditions for one month. The infection status was then assessed by RT-PCR on pools combining the RNA samples from four plants.

Results

A mixed viral infection on hollyhock

Symptoms consisting of vein chlorosis and yellow mosaic were noted in May 2021 on leaves of hollyhocks at two public sites in Lausanne and Nyon (Fig. 1a). Leaf samples from both locations were first checked for the presence of putative viral agents by TEM analysis of leaf dip preparations. Numerous isometric particles were visible in samples from both sites, while samples from Nyon contained additional flexuous particles > 800 nm in length (Fig. S1). These flexuous virions were reminiscent of members of *Potyvirus*, the largest genus of plant + ssRNA viruses, and this was confirmed by RT-PCR analyses using the generic primer pair ClFor/Rev (Table 1). Indeed, a cDNA of *ca.* 400 bp was amplified from RNA samples from Nyon but not from Lausanne. Blastn search on the sequenced amplicon revealed weak homology with the potyvirus Narcissus yellow stripe virus (GenBank acc. number LC314392.1, E-value = 6e-12, 80% nt identity), advocating that a novel potyviral species was present.

In order to gain insights into the isometric particles, a semi-purification protocol was performed on leaves collected in Lausanne. Two types of particles (*ca.* 30 nm in diameter) were clearly identified by TEM, corresponding presumably to intact virions and “empty” shells (Fig. 1b). Electrophoretic analysis of nucleic acid extracted from these particles disclosed a single RNA element that migrated slightly above a 3 kbp DNA marker (Fig. 1c, left). A protein of *ca.* 21 kDa was detected following particles denaturation and SDS-PAGE analysis (Fig. 1c, right). In parallel, a dsRNA element 6–7 kbp in length was extracted from symptomatic leaves from the same location (Fig. 1d). Altogether, these results were advocating for the presence of a tymovirus (1). This was tested by RT-PCR on RNA samples from both sites using the generic primer pair TymZ-F/R (Table 1). Both samples yielded a similar 340-bp amplicon that was highly similar to tymoviruses as revealed by Blastn search. The best hit (E-value = 3e-66, 80% nt identity) corresponded to the partially described Watercress white vein virus (WWVV, Genbank acc. number JQ001816.1, (48)), suggesting that a new tymovirus was present.

To investigate the pathogenicity of the novel tymovirus, leaf samples from both sites were used for mechanical inoculations of several indicator plants. Both inoculation procedures led to similar results. Plants of *Nicotiana occidentalis* showed stem curling and vein yellowing at 10 dpi (Fig. S2a-c). Plants of *N. benthamiana* turned yellow and exhibited mosaic symptoms at 12 dpi (Fig. S2d). Inoculated leaves of *Chenopodium quinoa* and *C. amaranticolor* showed local lesions after one week (Fig. S2e), but no systemic symptom was recorded on these species. Neither local nor systemic symptoms were visible on *N. clevelandii*, *N. tabacum* and *Brassica rapa* ssp. *chinensis*. Intriguingly, only RNA samples from *C. quinoa* and *C. amaranticolor* tested positive for the potyvirus by RT-PCR and no sample tested positive for

the tymovirus (Fig. S3). The symptoms observed on *N. occidentalis* and *N. benthamiana* were thus likely associated with another virus. TEM analysis of leaf dip preparations from these plants revealed flexuous particles *ca.* 600 nm in length (Fig. S4). RNA samples from these plants were thus tested by RT-PCR using the primer pair CarF2b/R targeting members of *Carlavirus* (Table 1), a large genus of flexuous + ssRNA viruses. The RT-PCR yielded an amplicon of *ca.* 750 bp, corresponding to a fragment of Gaillardia latent virus (GLV, Genbank acc. number LC485538.1, 96% nt identity, (49)). GLV was detected in all symptomatic indicator plants as well as the collected hollyhocks from both sites (Fig. S3).

Some tymoviruses have narrow host ranges limited to the family of their natural host (50, 51), and this trait could explain the lack of transmission of the newly identified tymovirus on indicator species. Semi-purified particles extracted from leaves collected in Lausanne were thus mechanically inoculated on the common mallow (*Malva sylvestris*), a fast-growing relative of hollyhock. No local symptom was observed but chlorotic spots and yellowing appeared at 12 dpi on upper leaves (Fig. 1e). RT-PCR analyses confirmed the presence of the tymovirus but not the potyvirus nor GLV (data not shown). This virus was thus hereafter named “Alcea yellow mosaic virus” (AYMV). Inoculations of *M. sylvestris* plants with GLV only (isolated on *N. occidentalis*) or GLV in combination with the potyvirus (isolated on *C. quinoa*) lead to symptomless infection and yellowing, respectively (data not shown).

Genomic characterization of AYMV

Given the symptoms observed on the common mallow, infected leaves of this plant were used for an extraction of viral dsRNA followed by HTS. Following Illumina sequencing and *de novo* assembly, numerous contigs showing significant similarities to tymoviruses were yielded (mean coverage = 357x, Fig. S5a). These contigs were then bridged by RT-PCR and Sanger sequencing. However, one region amplified with the primer pair TymoFinalF/R (Table 1) could neither be sequenced by Sanger sequencing nor cloned into the pGemT plasmid despite numerous attempts (data not shown). This amplicon was eventually sequenced by long read sequencing on a Nanopore Flongle flowcell (mean coverage = 80x, Fig. S5b). Following Sanger sequencing of the termini, a full-length genome was eventually compiled for AYMV and submitted to Genbank under accession number OP227146.

The length of AYMV gRNA is 6,444 nt, and its 5′ and 3′ untranslated regions (UTR) are 93 and 88 nt long, respectively (Fig. 2b). Tymoviral genomes were previously shown to be highly rich in cytidine (17), and this trait is exceptionally marked in the case of AYMV gRNA with 47.5% of C (Fig. 2a). In particular, the region that resisted Sanger sequencing and bacterial cloning (positions 1,826-2,644) is made of 58% of C with long repetitive stretches.

AYMV gRNA exhibits other features typical of tymoviruses: a protonable hairpin with internal C-A mismatches was identified at position 12–40 within the 5′-UTR (Fig. S6, $\Delta G = -6.20$ kcal/mol) and the tymobox was found starting at position 5,753, harboring one mismatch with the TYMV tymobox (Fig. 2c). AYMV 3′-UTR is likely to fold into a valine-accepting TLS given the extensive similarities with structural domains found in the 3′-UTRs of related viruses (Fig. 2d). The gRNA ends with a CCA motif which is

presumably involved in replication (17). Consistent with other tymoviruses, three ORFs encoding for MP, RP and CP were identified on AYMV genome (Fig. 2b, colored boxes).

AYMV MP is 702 AAs in length and shows low percentages of pairwise identities with other tymoviral MPs (30–43%, Fig. S7). A search in the CDD database failed to detect any conserved domain for this protein. It is particularly rich in proline (P, 25.9%) and leucine (L, 13.9%), and to a lesser extent in serine (S, 11.3%) and arginine (R, 9.0%) residues, and this AA bias is also observed for other tymoviral MPs (Fig. S8). In these proteins, the N-terminal moiety is more conserved than the C-terminal moiety, and this region harbors several conserved motifs and positively-charged AAs (Fig. S9-10). Unlike tymoviral CPs and RPs, MPs do not have direct homolog in *Marafivirus* and *Maculavirus*, the two other *Tymoviridae* genera. A few viruses of these genera harbor ORFs overlapping with RP, but no translation product has been evidenced and their role has not been established (53). In addition, these putative proteins do not contain the conserved motifs detected in tymoviral MPs and are much shorter (data not shown). Potential remote homologs of tymoviral MPs were searched using HHBlits (54), but only the aforementioned proteins of maculaviruses and marafiviruses were yielded, in line with the idea that tymoviral MPs are unique in the known global virome (55).

AYMV RP is 1894 AAs in length and shares 52–65% AA identities with other tymoviral RPs (Fig. S11, left). The following domains were predicted on this protein: Met pfam01660 (positions 39–349, E-value = $1.18\text{e-}74$), Pro cl05113 (positions 826–919, E-value = $1.21\text{e-}18$), Hel pfam01443 (positions 1,018 – 1,255, E-value = $2.50\text{e-}53$) and RdRp pfam00978 (position 1,480-1,827, E-value = 0). In addition, the domains Met and Pro are separated by a proline-rich region (PRR, with 24.4% of P) which corresponds to the region recalcitrant to Sanger sequencing and cloning. Based on previous studies (20), putative cleavage sites in the RP were identified at the motifs ⁹²⁴LGS/QP and ¹³⁰⁷VGS/SQ (Fig. 2b). AYMV CP is 189 AAs in length, shares 33–64% identities with other tymoviral CPs (Fig. S11, left) and harbors the tymovirus coat domain pfam00983 (E-value = $1.10\text{e-}35$) at positions 21–183.

Phylogenetic analyses

Among the three tymoviral proteins, the CP allows for the most inclusive phylogenetic analysis because 34 sequences for this protein have been deposited on NCBI (Table S1, excluding sequences from metagenomics data), while only 25 sequences are available for MPs and RPs. This is due to the use of generic primers amplifying the 3' region of gRNAs. A first ML phylogenetic tree was therefore constructed for an alignment of tymoviral CPs (Fig. 3, left). The topology of the obtained tree is in line with early serological studies that have evidenced a clear division of *Tymovirus* in two main clades (hereafter termed clade I and II), with the exception of the highly divergent *Erysimum latent virus* (ELV) (56). Clade I accommodates TYMV and other members mainly found on superrosids and for which fabaceous plants represent the largest group of natural hosts. Clade II contains viruses mainly retrieved from superrasterids, and in particular solanaceous plants. Members of clade I are grouped in two strongly supported subclades (IA and IB, 100% bootstrap replicates) and a third, less-supported subclade (IC, 68% bootstrap replicates). AYMV clusters with members of the latter which also gathers the other mallow-

infecting tymoviruses OMV and CYMV, four viruses found on *Fabaceae* as well as two viruses found on plants from the families *Passifloraceae* and *Apocynaceae*.

ML trees were then inferred for alignments of RPs and MPs (Fig. 3b-c). The topology of the RP tree is similar to the CP tree, exhibiting the same robust grouping into clades I and II. In this tree, AYMV is placed again within subclade IC, which is more strongly supported than in the CP tree. The MP tree differs by having members of clade II and subclade IC no longer hold into monophyletic groups. Nevertheless, AYMV is still most closely related to viruses from subclade IC in this tree.

Biological characterizations of AYMV

An experimental host range was assessed for AYMV based on sap inoculations. The virus was found to systemically infect all the tested mallows but not species from ten other families (Table S2). On susceptible hosts, mechanical inoculation was highly effective as all plants became symptomatic. Local lesions were rarely observed and symptoms typically appeared on upper leaves, consisting mostly in vein yellowing, mosaic, and chlorotic spots sometimes turning necrotic. The most severe symptoms were recorded on the common mallow and lavatera (*Lavatera trimestris*), on which most systematically infected leaves eventually became totally necrotic and fell off the plants. Notably, systemic infections were noted on the agronomically relevant mallows okra and cotton, on which pale green mottling were observed.

Specimens belonging to the weevil species *Aspidapion radiolus* and *Rhopalapion longirostre*, (*Coleoptera: Curculionoidea*), the mallow flea beetle *Podagrica fuscicornis* (*Coleoptera: Chrysomelidae*) were collected on healthy hollyhocks in a private garden in Prangins (district of Nyon, Switzerland) as they represent chewing insect that could serve as vector (Fig. 4a). In addition, specimens of the stink bug *Nezara viridula* (*Hemiptera: Pentatomidae*) were also collected from the same plants. These insects were used for transmission experiments, together with specimens of the aphid species *Myzus persicae* and *Aphis fabae* (*Hemiptera: Aphididae*). Even though all insects fed on plants, as evidenced by visible damages, transmission of AYMV was only observed with *A. radiolus* for which four plants became symptomatic and tested positive by RT-PCR among the 11 tested (Fig. 4b). In parallel, vertical transmissions was not recorded for 32 tested hollyhocks, suggesting that AYMV is not seed-borne.

Discussion

Here, classic diagnosis tools were used to untangle mixed infections in symptomatic hollyhocks collected at green sites from two Swiss cities, revealing the presence of three distinct + ssRNA viruses. Early results showed that AYMV was present at both sites and indicated that this virus represented a the first isolate of a probably new species, which prompted us to study this virus in particular.

We first aimed to obtain a full-length sequence for AYMV genome based on short-read sequencing of purified viral dsRNA together with bridging RT-PCR analyses. However, long-read sequencing was required to close the genome as a region could not be covered by short-read assembly and it was recalcitrant to

Sanger sequencing as well as bacterial cloning. The latter issue has been already described during the cloning of a similar region for the other tymovirus eggplant mosaic virus (57). This might have been caused by the high content in cytidine in the recalcitrant region, present as long repetitive stretches known to be associated with sequencing issues (58). The overall level of C is exceptionally high in AYMV gRNA, representing the record among tymoviruses. The reason for this strong nucleotide bias is unknown but could be related to efficient encapsidation. Indeed, under the light-induced, low-pH conditions found in the chloroplast-derived replication vesicles of TYMV, stretches of poly-Cs have been hypothesized to form tertiary structures required for virion assembly (21). The high C content is also directly related to the production of numerous P residues (encoded by CCC codons) in the MP and RP. It was recently shown that TYMV MP acts as a VSR by inhibiting the amplification step of the antiviral silencing, and co-localizes with siRNA bodies (59). One can speculate that the P residues in tymoviral MPs are involved, together with the conserved motifs and positively-charged AAs, in specific protein-protein or protein-RNA interactions within those bodies.

We identified other molecular and genomic features for AYMV that clearly associate this virus with tymoviruses, and phylogenetic trees further confirmed a robust positioning inside the genus. As illustrated here by the grouping of AYMV and other mallow-infecting members within subclade IC (Fig. 3), previous studies have highlighted the association between serological groups, natural hosts and host ranges among tymoviruses (56). However, despite several attempts, the molecular bases underlying this trait have not been found yet (57, 60).

The identification of a weevil vector for AYMV is rather interesting because, so far, *Scrophularia* mottle virus was the only member of the genus known to be transmitted by insects of this group (61). Most described vectors of tymoviruses are flea beetles from the genera *Epitrix* and *Diabrotica*, but other leaf-eating insects have been evidenced as vectors as well (24, 62, 63). Transmission assays have highlighted short acquisition time and long retention time, leading to the consensus that tymoviruses are likely “semi-persistently” transmitted in a mechanical process (13, 24). However, such a transmission mode still implies vector specificity, as noted here by the lack of transmission of AYMV using another weevil species and other chewing insects (although only a limited number of individuals were tested). In comparison to other insect groups, little is known regarding the spread of beetle vectors of plant viruses (64). It can be hypothesized that *A. radiolus* had been involved in the dissemination of AYMV in the two collection sites. Alternatively, given the high transmissibility of the virus, it is possible that it was carried on contaminated material via another unknown vector or through human activities.

In summary, in this study we have described the novel virus AYMV found in mixed infection on hollyhock in Switzerland. In the genus *Tymovirus*, the threshold for species demarcation is fixed at 90% AA identities for the CP by the International Committee on Taxonomy of Viruses, advocating that AYMV should be considered as the first isolate of a new viral species (1). Additional molecular and genomic analyses on GLV and the novel potyvirus will help better characterize these viruses as well as their interaction with AYMV. Since hollyhocks are widely distributed, these plants may serve as important reservoir for crop viruses, as noted by other studies (65, 66). In the case of AYMV, symptomatic infections

were evidenced on okra and cotton, advocating that this virus might represent a threat for these important crops should it be introduced in their respective cultivation areas.

Abbreviations

AA
amino acid
AYMV
Alcea yellow mosaic virus
CYMV
cacao yellow mosaic virus
CP
capsid protein
dsRNA
double-stranded RNA
dpi
days post inoculation
ELV
Erysimum latent virus
gRNA
genomic RNA
gRNA
genomic RNA
GLV
Gaillardia latent virus
Hel
helicase
HTS
high-throughput sequencing
ML
maximum likelihood
Met
methyltransferase
MP
movement protein
OMV
okra mosaic virus
Pro
papain-like protease
+ssRNA

positive-sense, single-stranded RNA
RP
replication protein
RdRp
RNA-dependent RNA polymerase
sgRNA
subgenomic RNA
TBMV
tomato blistering mosaic virus
TEM
transmission electron microscopy
tRNA
transfer RNA
TLS
tRNA-like structure
TYMV
turnip yellow mosaic virus
UTR
untranslated regions
VSR
viral suppressor of RNA silencing.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no external funding.

Authors' contributions

Designed the experiments and wrote the initial draft: MM. Performed the experiments: MM, JB, IK and AB. Collected the samples: MM, JB and IK. Analyzed the data: MM. Edited the manuscript: OS and AB. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank Marc Passerat, Sabine Bonnard and Larissa Grosu for plant maintenance and Floriane Bussereau and Julie Laurent for providing aphids. We are also grateful to Raphaël Groux for critical review of the manuscript.

References

1. Martelli GP, Sabanadzovic S, Abou-Ghanem Sabanadzovic N, Edwards MC, Dreher T. The family Tymoviridae. *Arch Virol* sept. 2002;147(9):1837–46.
2. Dreher TW, Goodwin JB. Transfer RNA mimicry among tymoviral genomic RNAs ranges from highly efficient to vestigial. *Nucleic Acids Research* 1 oct. 1998;26(19):4356–64.
3. Larson SB, Day J, Canady MA, Greenwood A, McPherson A. Refined structure of desmodium yellow mottle tymovirus at 2.7 Å resolution¹¹ Edited by T. Richmond. *J Mol Biology* 18 août. 2000;301(3):625–42.
4. Pflieger S, Blanchet S, Camborde L, Drugeon G, Rousseau A, Noizet M, et al. Efficient virus-induced gene silencing in Arabidopsis using a 'one-step' TYMV-derived vector. *Plant J.* 2008;56(4):678–90.
5. Powell JD, Barbar E, Dreher TW. Turnip yellow mosaic virus forms infectious particles without the native beta-annulus structure and flexible coat protein N-terminus. *Virol* 20 janv. 2012;422(2):165–73.
6. Kirino N, Inoue K, Tanina K, Yamazaki Y, Ohki ST. Turnip yellow mosaic virus isolated from Chinese cabbage in Japan. *J Gen Plant Pathol* 15 juill. 2008;74(4):331.
7. Stobbs LW, Cerkaskas RF, Lowery T, VanDriel L. Occurrence of Turnip Yellow Mosaic Virus on Oriental Cruciferous Vegetables in Southern Ontario, Canada. *Plant Disease* mars. 1998;82(3):351–1.
8. de Oliveira VC, Nagata T, Guimarães FC, Ferreira FA, Kitajima EW, Nicolini C, et al. Characterization of a novel tymovirus on tomato plants in Brazil. *Virus Genes* 1 févr. 2013;46(1):190–4.
9. Ferrand L, Nome C, Orílio AF, García ML, Nagata T, Ronco LB, et al. First Report of Tomato blistering mosaic virus Infecting Tomato in Argentina. *Plant Disease* mai. 2016;100(5):1026.
10. Gibbs A. A Plant Virus that Partially Protects Its Wild Legume Host against Herbivores. *INT.* 1980;13(1):42–7.
11. Min BE, Feldman TS, Ali A, Wiley G, Muthukumar V, Roe BA, et al. Molecular characterization, ecology, and epidemiology of a novel Tymovirus in *Asclepias viridis* from Oklahoma. *Phytopathol* févr. 2012;102(2):166–76.

12. Pagán I, Betancourt M, de Miguel J, Piñero D, Fraile A, García-Arenal F. Genomic and biological characterization of chiltepín yellow mosaic virus, a new tymovirus infecting *Capsicum annuum* var. *aviculare* in Mexico. *Arch Virol* mai. 2010;155(5):675–84.
13. Jones R, Fribourg C. CE. Beetle, contact and potato true seed transmission of Andean potato latent virus. *Ann Appl Biol*. 1977;86(1):123–8.
14. Mahgoub HA, Wipf-Scheibel C, Delécolle B, Pitrat M, Dafalla G, Lecoq H. Melon Rugose Mosaic Virus: Characterization of an Isolate from Sudan and Seed Transmission in Melon. *Plant Disease* 1 juin. 1997;81(6):656–60.
15. Koenig R, Barends S, Gultyaev AP, Lesemann DE, Vetten HJ, Loss S, et al. *Nemesia* ring necrosis virus: a new tymovirus with a genomic RNA having a histidylatable tobamovirus-like 3' end. *J Gen Virol* juin. 2005;86(Pt 6):1827–33.
16. Tzanetakis IE, Tsai CH, Martin RR, Dreher TW. A tymovirus with an atypical 3'-UTR illuminates the possibilities for 3'-UTR evolution. *Virol* 30 sept. 2009;392(2):238–45.
17. Dreher TW. Turnip yellow mosaic virus: transfer RNA mimicry, chloroplasts and a C-rich genome. *Mol Plant Pathol*. 2004;5(5):367–75.
18. Chen J, Li WX, Xie D, Peng JR, Ding SW. Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microRNA in host gene expression. *Plant Cell* mai. 2004;16(5):1302–13.
19. Ni F, Wu L, Wang Q, Hong J, Qi Y, Zhou X. Turnip Yellow Mosaic Virus P69 Interacts with and Suppresses GLK Transcription Factors to Cause Pale-Green Symptoms in *Arabidopsis*. *Mol Plant* 1 mai. 2017;10(5):764–6.
20. Jakubiec A, Drugeon G, Camborde L, Jupin I. Proteolytic Processing of Turnip Yellow Mosaic Virus Replication Proteins and Functional Impact on Infectivity. *J Virol* 15 oct. 2007;81(20):11402–12.
21. Rohozinski J, Hancock JM. Do light-induced pH changes within the chloroplast drive turnip yellow mosaic virus assembly? *J Gen Virol* févr. 1996;77(Pt 2):163–5.
22. Ding SW, Howe J, Keese P, Mackenzie A, Meek D, Osorio-Keese M, et al. The tymobox, a sequence shared by most tymoviruses: its use in molecular studies of tymoviruses. *Nucleic Acids Res* 11 mars. 1990;18(5):1181–7.
23. Brunt AA, Kenten RH, Gibbs AJ, Nixon HLY. Further Studies on Cocoa Yellow Mosaic Virus. *Microbiology*. 1965;38(1):81–90.
24. Givord L, Boer LD. Insect transmission of okra mosaic virus in the Ivory Coast. *Ann Appl Biol*. 1980;94(2):235–41.
25. Ndunguru J, Rajabu AC. Effect of okra mosaic virus disease on the above-ground morphological yield components of okra in Tanzania. *Scientia Horticulturae* 27 févr. 2004;99(3):225–35.
26. Reynard JS, Turco S, Brodard J, Kellenberger I, Maclot F, Schumpp O, et al. Identification and Molecular Characterization of a Novel Hordeivirus Associated With Yellow Mosaic Disease of Privet (*Ligustrum vulgare*) in Europe. *Front Microbiol* [Internet]. 2021 [cité 14 mai 2022];12. Disponible sur: <https://www.frontiersin.org/article/10.3389/fmicb.2021.723350>.

27. Mahillon M, Groux R, Bussereau F, Brodard J, Debonneville C, Demal S, et al. Virus Yellows and Syndrome “Basses Richesses” in Western Switzerland: A Dramatic 2020 Season Calls for Urgent Control Measures. *Pathogens* août. 2022;11(8):885.
28. Ha C, Coombs S, Revill PA, Harding RM, Vu M, Dale JL. Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. *Arch Virol*. 2008;153(1):25–36.
29. Sabanadzovic S, Ghanem-Sabanadzovic NA, Gorbalenya AE. Permutation of the active site of putative RNA-dependent RNA polymerase in a newly identified species of plant alpha-like virus. *Viol* 10 nov. 2009;394(1):1–7.
30. Nie X, Bai Y, Molen TA, Desjardins DC. Development of universal primers for detection of potato carlaviruses by RT-PCR. *J Virol Methods* mai. 2008;149(2):209–16.
31. Mahillon M, Decroës A, Peduzzi C, Romay G, Legrève A, Bragard C. 2021. RNA silencing machinery contributes to inability of BSBV to establish infection in *Nicotiana benthamiana*: evidence from characterization of agroinfectious clones of Beet soil-borne virus. *Journal of General Virology*. 2021;102(2):001530.
32. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinf* 1 août. 2014;30(15):2114–20.
33. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Curr Protocols Bioinf*. 2020;70(1):e102.
34. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* avr. 2012;9(4):357–9.
35. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinf* 15 sept. 2018;34(18):3094–100.
36. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 16 févr 2021;10(2):giab008.
37. Okonechnikov K, Golosova O, Fursov M, UGENE team. Unipro UGENE: a unified bioinformatics toolkit. *Bioinf* 15 avr. 2012;28(8):1166–7.
38. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinf* 1 mai. 2009;25(9):1189–91.
39. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 8 janv. 2020;48(D1):D265–8.
40. Mahillon M, Decroës A, Caulier S, Tiendrebeogo A, Legrève A, Bragard C. Genomic and biological characterization of a novel partitivirus infecting *Fusarium equiseti*. *Virus Res* 1 mai. 2021;297:198386.
41. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792–7.

42. Muhire BM, Varsani A, Martin DP. SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLOS ONE* 26 sept. 2014;9(9):e108277.
43. Thomsen MCF, Nielsen M. Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Research* 1 juill. 2012;40(W1):W281–7.
44. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* juin. 2017;14(6):587–9.
45. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* janv. 2015;32(1):268–74.
46. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol*. 2018;01(2):518–22. 35(.
47. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2 juill. 2019;47(W1):W256–9.
48. Harju V, Skelton A, Forde S, Bennett S, Glover R, Monger W, et al. New virus detected on *Nasturtium officinale* (watercress). *New Disease Reports*. 2012;25(1):1–1.
49. Menzel W, Hamacher J, Winter S. CHARACTERIZATION OF A. NEW CARLAVIRUS FROM GAILLARDIA ARISTATA. *Acta Hortic.* févr 2015;(1072):129–33.
50. Gibbs AJ, Mackenzie AM, Abdul-Samad N. A tymovirus from *Calopogonium mucunoides* in Malaysia is not clitoria yellow vein tymovirus. *Arch Virol*. 1997;142(8):1697–702.
51. Nicolini C, Pio-Ribeiro G, Andrade GP, Melo FL, Oliveira VC, Guimarães FC, et al. A distinct tymovirus infecting *Cassia hofmannseggii* in Brazil. *Virus Genes* 1 août. 2012;45(1):190–4.
52. de Smit MH, Gultyaev AP, Hilge M, Bink HHJ, Barends S, Kraal B, et al. Structural variation and functional importance of a D-loop–T-loop interaction in valine-accepting tRNA-like structures of plant viral RNAs. *Nucleic Acids Research* 1 oct. 2002;30(19):4232–40.
53. Edwards MC, Weiland JJ, Todd J, Stewart LR, Lu S. ORF43 of maize rayado fino virus is dispensable for systemic infection of maize and transmission by leafhoppers. *Virus Genes* 1 avr. 2016;52(2):303–7.
54. Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods* févr. 2012;9(2):173–5.
55. Dolja VV, Krupovic M, Koonin EV. Deep Roots and Splendid Boughs of the Global Plant Virome. *Annu Rev Phytopathol* 25 août. 2020;58:23–53.
56. Guy PL, Dale JL, Adena MA, Gibbs AJ. A taxonomic study of the host ranges of tymoviruses. *Plant Pathol*. 1984;33(3):337–46.
57. Skotnicki ML, Ding SW, Mackenzie AM, Gibbs AJ. Infectious eggplant mosaic tymovirus and ononis yellow mosaic tymovirus from cloned cDNA. *Archives of Virology* 1 mars. 1993;131(1):47–60.

58. Kieleczawa J. Fundamentals of sequencing of difficult templates—an overview. *J Biomol Tech* juill. 2006;17(3):207–17.
59. Sehki H, Yu A, Elmayan T, Vaucheret H. TYMV has dual action on the plant RNA silencing defense through its VSR P69 and the host RNASE THREE LIKE1 [Internet]. *bioRxiv*; 2022 [cité 8 oct 2022]. p. 2022.03.30.486349. Disponible sur: <https://www.biorxiv.org/content/10.1101/2022.03.30.486349v1>.
60. Srifah P, Keese P, Weiller G, Gibbs A. Comparisons of the genomic sequences of *erysimum latent virus* and other tymoviruses: a search for the molecular basis of their host specificities. *J Gen Virol* juin. 1992;73(Pt 6):1437–47.
61. Weidemann HL. Zur Übertragung des *Scrophularia Mottle Virus*. *J Phytopathol* nov. 1973;78(3):278–81.
62. Guy PL, Gibbs AJ. Further studies on turnip yellow mosaic tymovirus isolates from an endemic Australian *Cardamine*. *Plant Pathol*. 1985;34(4):532–44.
63. Markham R, Smith KM. Studies on the virus of turnip yellow mosaic. *Parasitol* févr. 1949;39(3–4):330–42.
64. Reynolds DR, Chapman JW, Harrington R. The Migration of Insect Vectors of Plant and Animal Viruses. In: *Advances in Virus Research* [Internet]. Academic Press; 2006 [cité 26 sept 2022]. p. 453–517. (Plant Virus Epidemiology; vol. 67). Disponible sur: <https://www.sciencedirect.com/science/article/pii/S0065352706670127>.
65. Khan A, Imtiaz Y, Tahir M, Briddon RW. Complete genome sequence of hollyhock vein yellowing virus, a novel monopartite begomovirus infecting hollyhock in Pakistan. *Arch Virol* sept. 2021;166(9):2607–10.
66. Park WM, Park SK, Yoon JY, Ryu KH, Park JK. First Report of Zucchini yellow mosaic virus on Hollyhock (*Althaea rosea*). *Plant Pathol J*. 2002;18(3):121–5.

Figures

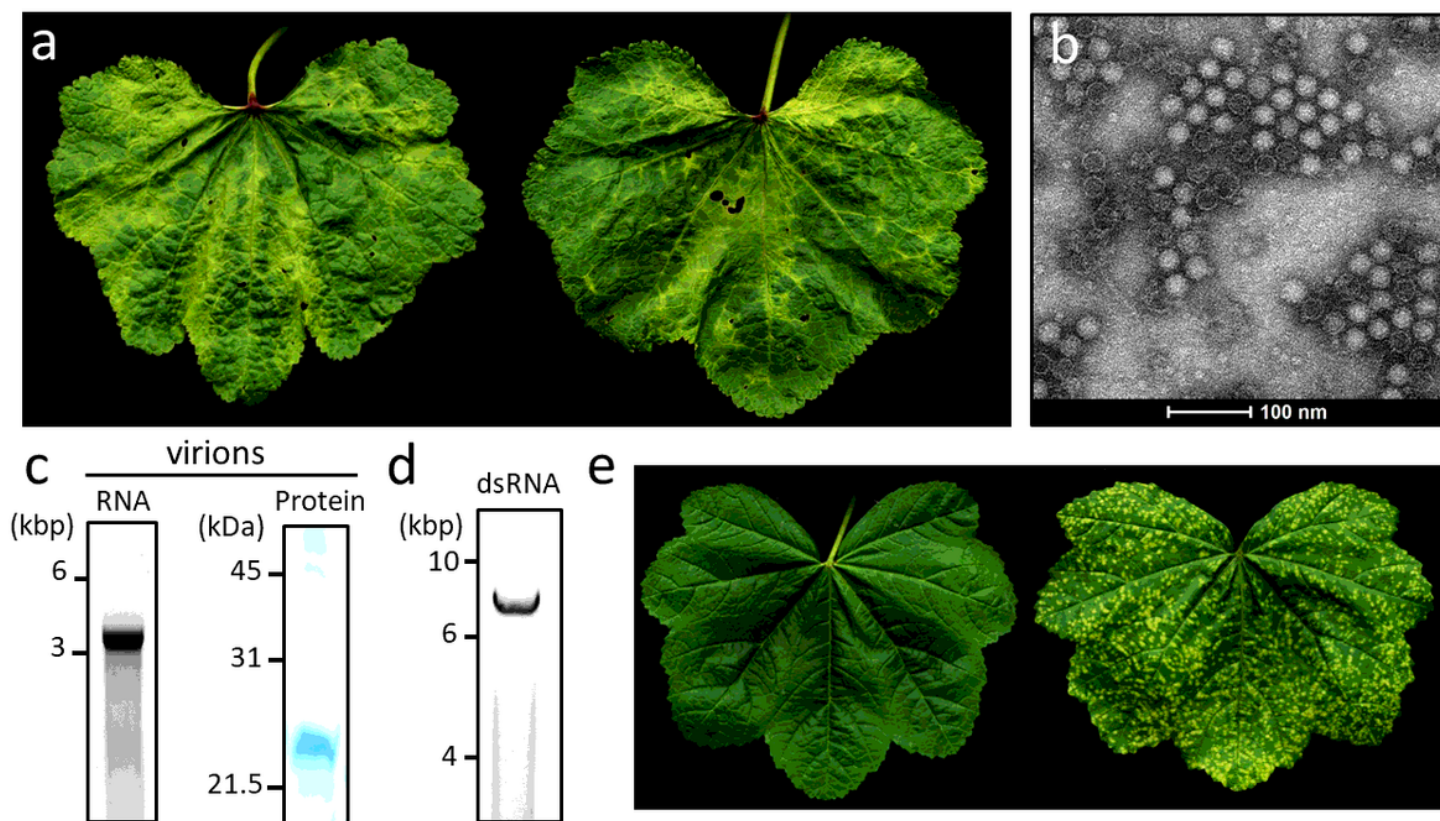


Figure 1

Identification of the novel tymovirus AYMV. **(a)** Hollyhock leaves collected in Lausanne showing vein clearing, chlorosis and mosaic symptoms. **(b)** Transmission electron micrograph of semi-purified virions negatively stained with 4% phosphotungstate. Note the staining penetration in empty particles. **(c)** Electrophoresis gel for RNA (left) and protein (right) associated with semi-purified particles. **(d)** Electrophoresis gel for the dsRNA extracted from symptomatic hollyhock leaves. **(e)** Healthy leaf of *M. sylvestris* (left) in comparison with an AYMV-infected leaf showing chlorotic spots (right), at 30 dpi.

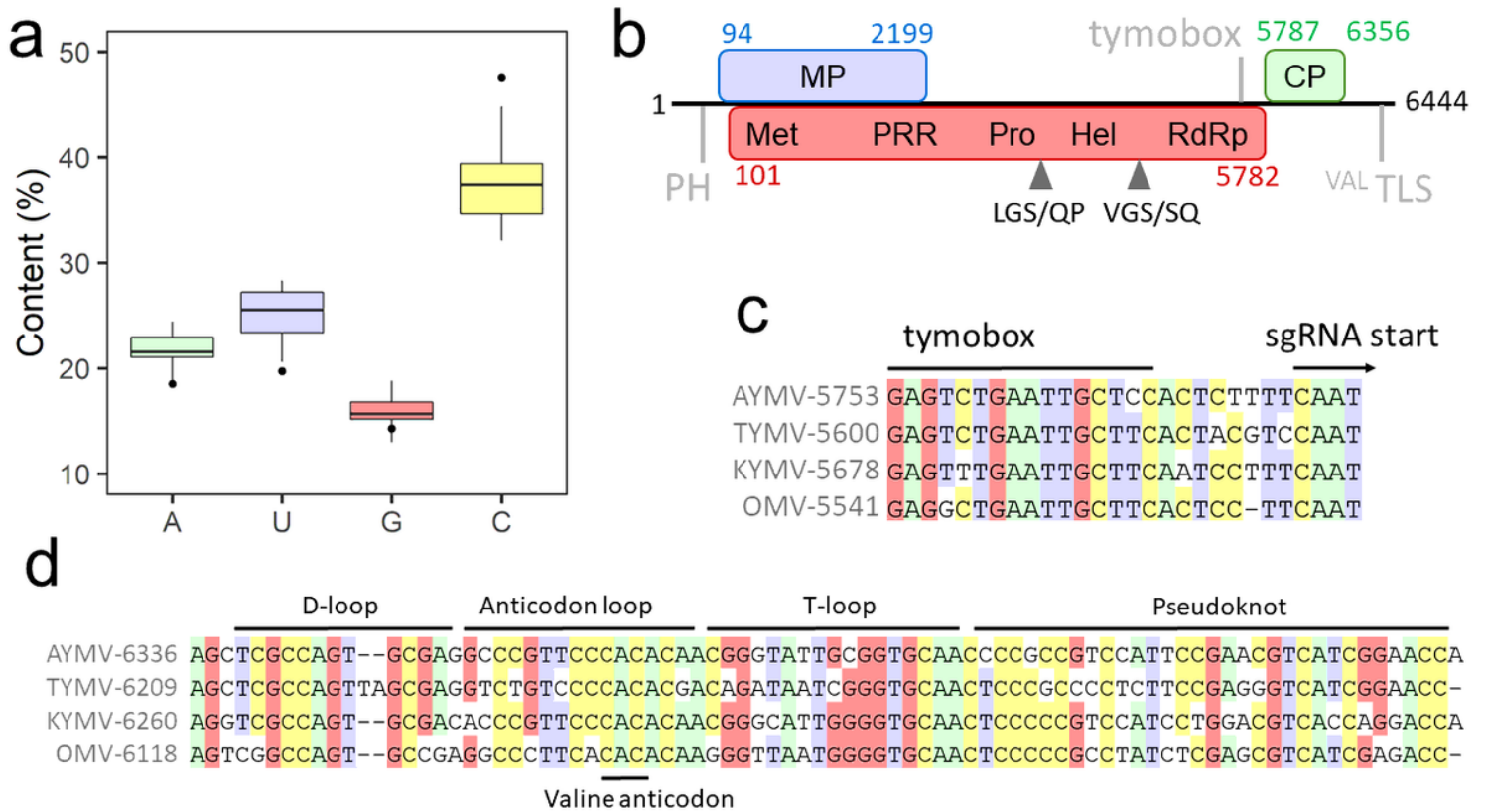


Figure 2

Characterization of AYMV genome. **(a)** Base content for the gRNA of tymoviruses (box plots) and AYMV (black points). **(b)** Schematic representation of AYMV gRNA. ORFs are represented as colored boxes with associated positions and conserved domains/functions (see main text for details). The putative RP cleavage sites are indicated by gray arrows. The protonable hairpin (PH), tymobox and valine tRNA-like structure (^{Val}TLS) are indicated by vertical gray lines. **(c)** Alignment of the tymobox cDNA for AYMV and selected related species. **(d)** Alignment of the 3'-UTR cDNA for the same viruses described in C. The structural annotations are reproduced from a previous analysis of tymoviral tRNA (52). Colored background indicate $\geq 75\%$ conservation. Full virus names and accession numbers are listed in Table S1.

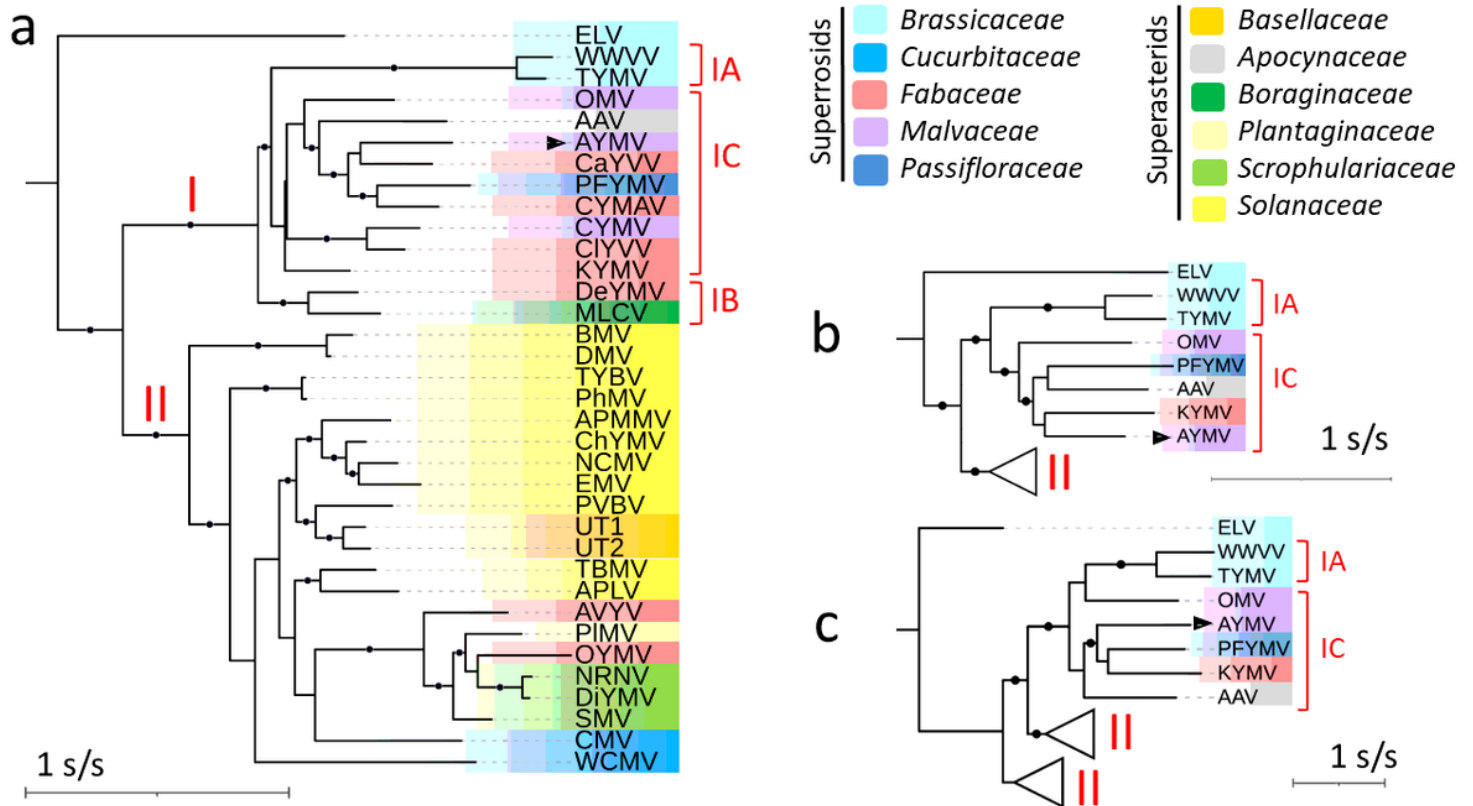


Figure 3

ML phylogenetic tree of the tymoviral proteins. **(a)** CP tree (model: rtREV+F+I+G4). **(b)** RP tree (model: rtREV+F+I+G4). **(c)** MP tree (model cpREV+F+I+G4). Full virus names and accession numbers are listed in Table S1. The position of AYMV is highlighted by a black arrow. Selected members of the genera *Marafivirus* and *Maculavirus* were used to root the initial CP and RP trees. Black circles on branches indicate bootstrap support >70%. The tree scales are given in substitution per site. Colored background are indicative of the family of the natural host.

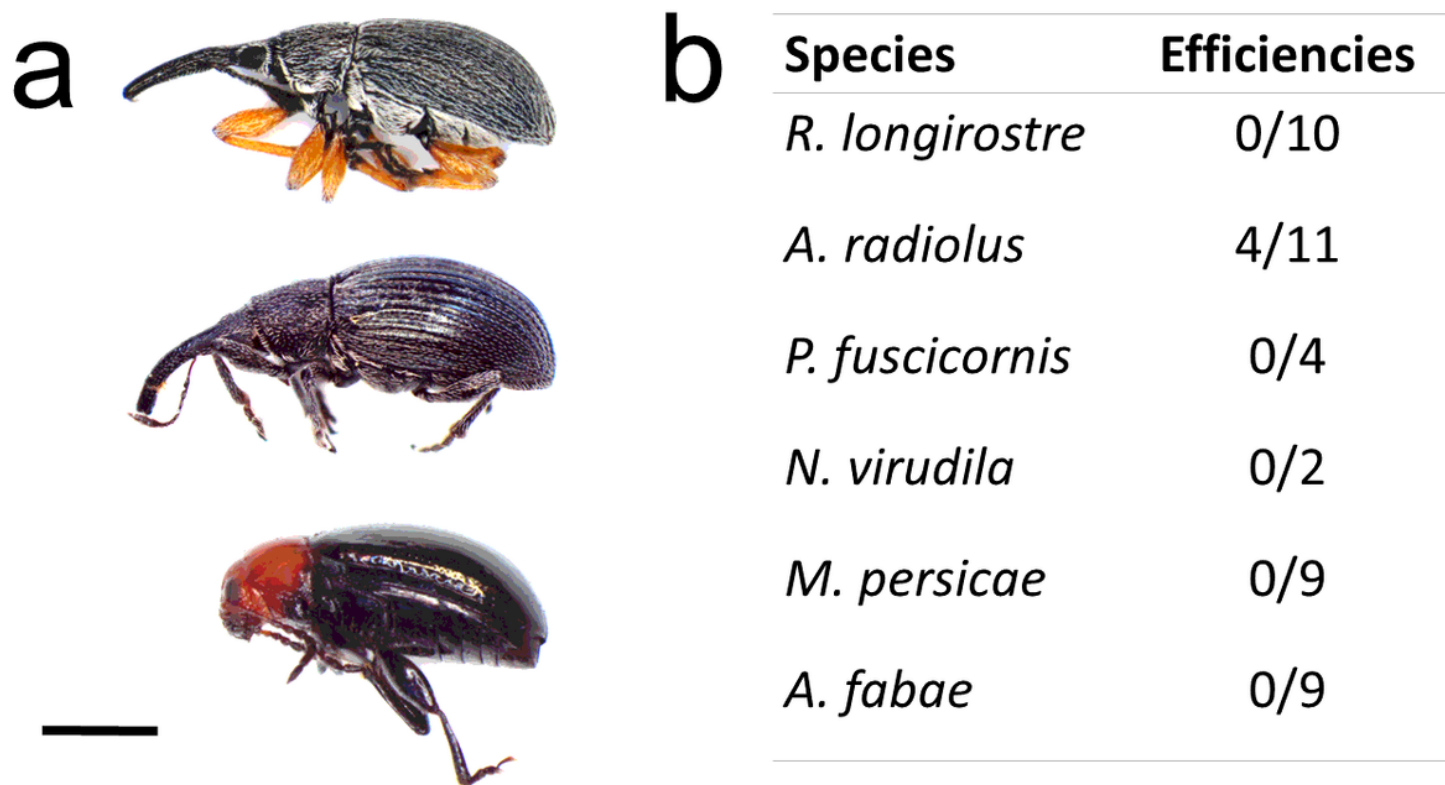


Figure 4

Insect transmission assays. **(a)** Specimens of *Coleoptera* found on hollyhock and used for transmissions assay. Up: *R. longirostre*, middle: *A. radiolus* and low: *P. fuscicornis*. The bar represents 1 cm. **(b)** Transmission efficiencies for the different insect species. Pairs of insects were used for each transmission assay.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigures.pdf](#)
- [SupplementaryTables.pdf](#)