Virus Evolution, 2024, **10(1)**, veae025
DOI: https://doi.org/10.1093/ve/veae025
Advance Access Publication 9 March 2024
Research Article

Virome release of an invasive exotic plant species in southern France

Oumaima Moubset,[1,2] Denis Filloux,[1,2] Hugo Fontes,[3,4,‡] Charlotte Julian,[1,2] Emmanuel Fernandez,[1,2] Serge Galzi,[1,2] Laurence Blondin,[1,2] Sélim Ben Chehida,[5] Jean-Michel Lett,[5] François Mesléard,[3,4] Simona Kraberger,[6] Joy M. Custer,[6] Andrew Salywon,[7] Elizabeth Makings,[8] Armelle Marais,[9] Frédéric Chiroleu,[5] Pierre Lefeuvre,[5,§] Darren P. Martin,[10,**] Thierry Candresse,[9,††] Arvind Varsani,[6,11,‡‡] Virginie Ravigné,[1,2,†] and Philippe Roumagnac[1,2,†,§§,*]

[1]UMR PHIM, CIRAD, Baillarguet TA A-54/K, Montpellier 34090, France, [2]PHIM Plant Health Institute, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, Baillarguet TA A-54/K, Montpellier 34090, France, [3]Tour du Valat, Institut de recherche pour la conservation des zones humides méditerranéennes, Le Sambuc, Arles 13200, France, [4]Institut Méditerranéen de Biodiversité et Ecologie, UMR CNRS-IRD, Avignon Université, Aix-Marseille Université, IUT d'Avignon, Avignon 84911, France, [5]UMR PVBMT, CIRAD, Saint-Pierre, La Réunion F-97410, France, [6]The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA, [7]Department of Research, Conservation and Collections, Desert Botanical Garden, Phoenix, AZ 85008, USA, [8]Vascular Plant Herbarium, School of Life Sciences, Arizona State University, 734 West Alameda Drive, Tempe Tempe, AZ 85282, USA, [9]UMR BFP, University Bordeaux, INRAE, Villenave d'Ornon 33140, France, [10]Division of Computational Biology, Department of Integrative Biomedical Sciences, Institute of infectious Diseases and Molecular Medicine, University of Cape Town, Anzio Rd, Cape Town 7925, South Africa and [11]Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Observatory, Cape Town 7700, South Africa
[†]These authors contributed equally to this work.
[‡]https://orcid.org/0000-0002-7826-6680
[§]https://orcid.org/0000-0003-2645-8098
[**]https://orcid.org/0000-0002-8785-0870
[††]https://orcid.org/0000-0001-9757-1835
[‡‡]https://orcid.org/0000-0003-4111-2415
[§§]https://orcid.org/0000-0001-5002-6039
*Corresponding author: E-mail: philippe.roumagnac@cirad.fr

Abstract

The increase in human-mediated introduction of plant species to new regions has resulted in a rise of invasive exotic plant species (IEPS) that has had significant effects on biodiversity and ecosystem processes. One commonly accepted mechanism of invasions is that proposed by the enemy release hypothesis (ERH), which states that IEPS free from their native herbivores and natural enemies in new environments can outcompete indigenous species and become invasive. We here propose the virome release hypothesis (VRH) as a virus-centered variant of the conventional ERH that is only focused on enemies. The VRH predicts that vertically transmitted plant-associated viruses (PAV, encompassing phytoviruses and mycoviruses) should be co-introduced during the dissemination of the IEPS, while horizontally transmitted PAV of IEPS should be left behind or should not be locally transmitted in the introduced area due to a maladaptation of local vectors. To document the VRH, virome richness and composition as well as PAV prevalence, co-infection, host range, and transmission modes were compared between indigenous plant species and an invasive grass, cane bluestem (*Bothriochloa barbinodis*), in both its introduced range (southern France) and one area of its native range (Sonoran Desert, Arizona, USA). Contrary to the VRH, we show that invasive populations of *B. barbinodis* in France were not associated with a lower PAV prevalence or richness than native populations of *B. barbinodis* from the USA. However, comparison of virome compositions and network analyses further revealed more diverse and complex plant–virus interactions in the French ecosystem, with a significant richness of mycoviruses. Setting mycoviruses apart, only one putatively vertically transmitted phytovirus (belonging to the *Amalgaviridae* family) and one putatively horizontally transmitted phytovirus (belonging to the *Geminiviridae* family) were identified from *B. barbinodis* plants in the introduced area. Collectively, these characteristics of the *B. barbinodis*-associated PAV community in southern France suggest that a virome release phase may have immediately followed the introduction of *B. barbinodis* to France in the 1960s or 1970s, and that, since then, the invasive populations of this IEPS have already transitioned out of this virome release phase, and have started interacting with several local mycoviruses and a few local plant viruses.

Keywords: viral ecology; viral metagenomics; plant-associated viruses; invasive exotic plant species.

Introduction

Over recent decades, anthropogenic activities, ranging from unintentional plant transfers via global trade to purposeful plant introductions for agriculture or ornamental reasons, have facilitated the spread of invasive exotic plant species (IEPS) (Hulme 2009; Seebens et al. 2017), leading to a marked increase in their

distributions and abundances (Pyšek et al. 2020). IEPS pose severe threats to global biodiversity and ecosystem stability (Pyšek and Richardson 2010; Vilà et al. 2010; Hood-Nowotny et al. 2023) by competing with native species (Gaertner et al. 2009) in ways that disturb ecosystem processes and functionalities (Bartz and Kowarik 2019; Rai and Singh 2020).

Among the most commonly cited explanations for the global success of IEPS is the enemy release hypothesis (ERH) (Mack et al. 2000; Maron and Vilà 2001), which postulates that IEPS free from their native herbivores and natural enemies in new environments can outcompete indigenous species and become invasive (Keane and Crawley 2002). Specifically, the ERH predicts that in the introduced range: (1) the specialist enemies of IEPS will rarely be present (even if a few enemies are able to survive on the invasive propagules), (2) host switching of native species-associated specialist enemies will be infrequent, and (3) the generalist enemies will have a greater impact on native plant species (Keane and Crawley 2002). However, evidence of absence of enemies in the introduced range does not necessarily mean that the ERH is supported. Two other conditions should be fulfilled, first that enemies are important regulators of IEPS fitness and second that IEPS can capitalize on reduced enemy pressure, resulting in increased competitive ability and population growth (Brian and Catford 2023). In addition, the effect of the enemy release should be modulated by the biotic and abiotic contexts, including e.g. resource availability, phylogenetic relatedness of exotic and native species, or the time since introduction of an IEPS (Brian and Catford 2023).

Hence, enemy release is generally considered as a temporary advantage for the IEPS in the environment it is introduced to: an advantage that should fade as the IEPS colonizes the environment and becomes a target of local enemies (Flory, Clay, and Thrall 2013; Siemann, Rogers, and Dewalt 2006; Mitchell et al. 2010; Li et al. 2022). While newly introduced plants are hypothesized to initially exhibit benefits under the ERH, they tend to gradually accumulate novel enemies over time periods ranging from 50 to 200 years post-introduction (Hawkes 2007). Echoing this, Mitchell et al. (2010) showed that IEPS introduced 400 years ago supported six times more pathogen species than those introduced 40 years ago. Taken together, these studies suggest that the benefits to IEPS predicted by the ERH could be transient, diminishing over decades-long timeframes (Brian and Catford 2023).

Approximately 60 per cent of studies of vertebrate-, invertebrate-, or plant-interactions with predators, parasites, herbivores, and pathogens found that invasive populations were less prone to fitness costs attributable to these enemies than native populations: an observation congruent with the foundation of the ERH (Colautti et al. 2004). Many of these studies have, for example, reported that the richness of herbivorous insect communities is greater on native plants than on introduced plants, as herbivorous insects tend to favor native species over introduced species (Liu and Stiling 2006; Meijer et al. 2016). Similarly, Mitchell and Power, (2003) estimated the richness of viruses and fungi that infected 473 plant species introduced from Europe to the USA. They found that relative to their native range, plants in their introduced range were on average infected by 84 per cent fewer rust, smut, and powdery mildew fungi and 24 per cent fewer viruses. Since these seminal studies, only a few others have examined virus communities in IEPS (Rua et al. 2011). As a complement to the Mitchell and Power (2003) analysis, two metagenomics-based studies have explored differences in viral prevalence between exotic and indigenous plants. First, a study focusing on grass-infecting mastreviruses on Réunion Island in the South-West Indian Ocean reported that indigenous plant species tended to show higher viral prevalence than IEPS (Claverie et al. 2023). In contrast, a study using a spatial metagenomics approach revealed that IEPS in South Africa had a greater prevalence of PAV families than indigenous plants (Bernardo et al. 2018).

In addition to IEPS benefitting from reduced herbivory after introduction, they are also likely to benefit from both an absence of harmful microorganisms and the presence of beneficial microorganisms that could bolster their adaptive advantages (Faillace, Lorusso, and Duffy 2017). For instance, many IEPS are introduced into new environments via seeds, which can harbor endophytes—organisms that live within plants without causing apparent harm (Lehan et al. 2013). In many cases, these endophytes confer various benefits to their host plants, such as enhanced nutrient uptake, resistance to pathogens or predators, and increased tolerance to environmental stresses (Elsheikh et al. 2021; Geisen et al. 2021; Jeong et al. 2021; War et al. 2023). As IEPS grow in their new environments, the presence of such beneficial endophytes could further amplify the advantages predicted by the ERH, enhancing the chances of IEPS survival, spread, and establishment in the introduced range.

The vast majority of described viruses have consequences on their host fitness ranging from negligible to strongly deleterious, so that losing viruses would on average mechanically involve a fitness gain. However, recently, it has been proposed that some viruses could be beneficial to their host (Roossinck and Schultz-Cherry 2015a). These mutualist interactions could be mainly expected among viruses that have a long history of association with their host plant, and thus most likely among plant viruses that are vertically transmitted (Roossinck 2011). Few studies have characterized the diversity and abundance of vertically transmitted plant viruses in natural settings and evidence of mutualistic plant viruses has been sparse (Roossinck 2015b). Yet, the possibility that some viruses could be beneficial to their host has to be accounted for to infer the possible consequences of virus loss on IEPS fitness. We therefore propose the virome release hypothesis (VRH) as a virus-centered hypothesis inspired from the original ERH. The VRH predicts that (1) horizontally transmitted viruses that associated with IEPS in their native geographical range would often be left behind during the introduction process or could not be locally transmitted in the introduced range due to maladaptation of local vectors and (2) vertically transmitted PAV would be co-introduced during the dissemination of the IEPS. While about 25 per cent of all known PAV can potentially be vertically transmitted through seeds (Gutiérrez-Sánchez et al. 2023), the vast majority of PAV are horizontally transmitted, meaning that a measurable prediction of the VRH is that IEPS should have a reduced virome richness in the introduced range.

Here, we tested the VRH focusing on the cane bluestem (*Bothriochloa barbinodis* (Lag.) Herter, Poaceae), a perennial warm-season C4-grass species native to the Americas, which ranges from Bolivia and Argentina to the southwestern USA (Vega 2000). Accidentally introduced to France via the wool trade between 1964 and 1976, this species first became established in fallow lands around the French Mediterranean coast and subsequently spread throughout France along road and rail routes (Fried 2017). Specifically, we compared virome richness and composition as well as PAV prevalence, co-infection, host range, and transmission modes between indigenous grass species and *B.*

barbinodis sampled from both one area of the native range of *B. barbinodis* in the USA (Arizona) and its introduced range in southern France.

We found that, contrary to the VRH, invasive populations of *B. barbinodis* in France were not associated with a lower viral prevalence or richness than native populations of *B. barbinodis* from the USA. However, while *B. barbinodis* had one of the highest viral richness in the native area compared to other native grasses, it had one of the lowest in the invaded area. Additionally, comparison of virome compositions and network analyses further revealed significantly higher richness and more frequent horizontal transmission of mycoviruses in France than in the Arizona desert. Setting these mycoviruses apart, only two PAV infected *B. barbinodis* in France, including one putatively vertically transmitted PAV (belonging to the *Amalgaviridae* family) and one putatively horizontally transmitted PAV (belonging to the *Geminiviridae* family). Collectively, these characteristics of the *B. barbinodis*-associated PAV community in southern France suggest that a virome release phase may have immediately followed the introduction of *B. barbinodis* to France in the 1960s or 1970s, and that, since then, the invasive populations of this IEPS have already transitioned out of this virome release phase, and have started interacting with several local mycoviruses and a few local plant viruses.

Materials and methods

Bothriochloa barbinodis invasion of France and field surveys

Bothriochloa barbinodis has been classified as invasive by both the National Botanic Conservatories of Méditerranée and Midi-Pyrénées which are in charge of the surveillance of southern France (Cottaz et al. 2021). *B. barbinodis*, which was initially restricted to the Lower Languedoc region of France for about 40 years, has now invaded two-thirds of the country (Supplementary Figure S1A). Two sampling surveys were conducted in France (*B. barbinodis* introduced area) in July 2018 and the USA (*B. barbinodis* native area) in October 2021. A total of 521 plant samples, including both indigenous grass species and *B. barbinodis*, were collected from 14 sites: 320 samples from eight sites (sites 1 to 8) in southern France and 201 samples from six sites (sites 9 to 14) in Arizona, USA (Table 1, Supplementary Figure S1 and Supplementary Table S1). While the sampling sites in Arizona were located in natural areas of the Sonoran Desert, the sampling sites in France were fallow lands at the edges of cropping areas. Four and nine indigenous perennial grass species were, respectively, collected in France (*Brachypodium phoenicoides* (L.) Roem. & Schult., *Cynodon dactylon* (L.) Pers., *Dactylis glomerata* L. and *Elytrigia campestris* (Godr & Gren.) Kerguélen) and Arizona (*Aristida purpurea* Nutt., *Bouteloua aristidoides* (Kunth) Griseb., *Bouteloua curtipendula* (Michx.) Torr., *Bouteloua hirsuta* Lag., *Cynodon dactylon, Heteropogon contortus* (L.) P. Beauv., *Hilaria belangeri* (Steud.) Nash, *Panicum hirticaule* J. Presl and *Sorghum halepense* Pers.). A total of 10 *B. barbinodis* individuals and 10 individuals from three indigenous perennial grass species were collected without regard to symptoms from each site in France, resulting in a total of 40 samples per site. In Arizona, 8 to 11 individuals of *B. barbinodis* and 5 to 10 individuals of two to three indigenous perennial grass species were collected without regard to symptoms from each site (Supplementary Table S1). All samples were immediately refrigerated at 4°C in the field and transported with 4°C refrigeration to laboratories in either Phoenix (in the USA) or Montpellier (in France). The French samples were further kept at −80°C before being processed. The USA samples were dried at 50°C for 48 h and sent to Montpellier for further processing.

Virion-associated nucleic-acid-based viral metagenomics

The 521 collected plant samples were individually processed using the virion-associated nucleic acid (VANA)-based metagenomics approach (François et al. 2018). This approach is useful for inventorying plant viromes, as it enables the detection of both DNA and RNA viruses from plant samples that have been preserved for storage using either freezing or dehydration (Moubset et al. 2022; Schönegger et al. 2023). The VANA approach is based on a primer tagging strategy using 24 nt long multiplex identifier (MID) that enables the tracing of each sequencing read in pooled samples back to the georeferenced plant sample from which it came (Roossinck et al. 2010). One to three negative controls and one positive control were used in each of the five libraries, each containing 96 samples from the sequencing run that included the French dataset, resulting in a total of 15 controls. The positive control was an alfalfa (*Medicago sativa L.*) plant maintained at Cirad (Montpellier) co-infected by alfalfa mosaic virus (*Alfamovirus* genus, *Bromoviridae* family) and bean leafroll virus (*Luteovirus* genus, *Tombusviridae* family). The negative control was an uninfected sugarcane seedling, obtained through strict quarantine procedures in our laboratory at Cirad (Montpellier). The second sequencing run, containing the American dataset, consisted of four libraries with 96 samples each. This run included the same negative control used in the French dataset run and two positive controls, amounting to a total of 12 controls. The positive controls comprised one sugarcane sample co-infected by sugarcane streak Egypt virus (*Mastrevirus* genus, *Geminiviridae* family) and sugarcane white streak virus (*Mastrevirus* genus, *Geminiviridae* family), and one sugarcane sample infected by sugarcane yellow leaf virus (*Polerovirus* genus, *Solemoviridae* family). In addition to the 521 samples from the current study and the 27 control samples, 316 plant samples from other metagenomics projects conducted by the Montpellier research group complemented the 9 libraries. Amplicons obtained following the VANA metagenomics process were sequenced by Azenta (Leipzig, Germany) on an Illumina HiSeq 3000 sequencer (2 × 150 nt sequencing). As previously described by François et al. (2018), the short sequence reads obtained were cleaned and assembled into longer continuous sequences (contigs) using the SPAdes Genome Assembler 3.6.2 (Bankevich et al. 2012). Taxonomic assignment of the cleaned Illumina reads and contigs was carried out using the BLASTx-equivalent algorithm implemented in DIAMOND v0.9.22 (Buchfink, Xie, and Huson 2015) against the NCBI non-redundant (nr) protein database with an e-value threshold of <0.001. On the basis of the best hit returned by DIAMOND, related groups of virus-like sequences (operational taxonomic units, OTUs) were assigned to known plant virus genera and to a tentative species. The cleaned reads have been deposited in the sequence read archive of GenBank (accession number PRJNA721112 for the French plant samples and accession number PRJNA1020112 for the American plant samples).

Setting nucleotide sequence length thresholds

The VANA Illumina approach produced sequences of variable lengths, sequences ≤126 nt (i.e. 150 nt long reads generated by Illumina HiSeq less the 24 nt long multiplex identifier) generally being individual sequencing reads and sequences ≥126 nt being contigs assembled from two or more sequencing reads. A recent simulation study of VANA-produced sequences revealed that the lengths

Table 1. Data summary of plant sampling surveys, yields of VANA Illumina reads, observed virus richness, observed virus prevalence, and transmission modes of plant associated viruses (PAV) including phytoviruses (PHY) and mycoviruses (MYC) infecting indigenous Poaceae and *Bothriochloa barbinodis* in both French and American ecosystems

	Total plant species		Indigenous poaceae [a]		*Bothriochloa barbinodis*	
	France	USA	France	USA	France	USA
Number of sites	8	6	8	6	8	6
Number of plant species	5	10	4	9	1	1
Number of collected individuals	320	201	240	143	80	58
Number of good reads	71,106,616	82,327,473	53,861,735	62,921,097	17,244,881	19,406,376
Mean number of good reads per plant sample (SD)	222,208 (168,887)	409,589 (440,322)	224,424 (169,442)	440,008 (441,448)	215,561 (168,097)	334,593 (432,166)
Number of PAV reads (% of good reads)	4,987,513 (7%)	885,144 (1%)	4,181,524 (7.7%)	824,188 (1.3%)	805,989 (4.6%)	60,956 (0.3%)
Mean number of PAV reads per plant sample (SD)	15,586 (48,504)	4,404 (19,327)	17,423 (49,585)	5,764 (22,723)	10,075 (44,952)	1,051 (2,877)
Number of contigs assigned to viral OTUs	44,047	4,097	42,009	3,362	2,038	735
Mean sequence length of contigs assigned to viral OTUs (SD)	260 (152)	251 (693)	270 (157)	275 (820)	228 (130)	192 (71)
Mean number of contigs assigned to viral OTUs per plant sample (SD)	138 (1,587)	20 (34)	175 (1,831)	24 (40)	26 (37)	13 (10)
Number of PAV species (%)	49 (100%)	26 (100%)	43	16	12	12
Number of PHY species (%)	29 (59.1%)	19 (73.0 %)	25	10	4	9
Number of MYC species (%)	20 (40.8%)	7 (26.9%)	18	6	8	3
% infected plants (PAV) (SD)	21.5 (4)	9.9 (6)	23.7 (5)	5.6 (8)	15 (10)	20.7 (11)
% infected plants (PHY) (SD)	8.8 (5)	7.9 (6)	10.4 (6)	4.1 (8)	3.7 (10)	17.2 (11)
% infected plants (MYC) (SD)	14.4 (5)	2.4 (6)	15 (5)	2 (8)	12.5 (10)	3.4 (12)
Number of vertically transmitted virus species	13	11	8	8	5	5
Number of horizontally transmitted virus species	26	15	23	8	3	7

[a] *Bothriochloa barbinodis* was excluded from Indigenous Poaceae in the USA; (%) infected plants refers to the observed prevalence (the percentage of individuals within a specific population that are infected with at least one PAV, PHY, or MYC).
PS: For a better visualization of these data, check the Venn diagrams (Supplementary Figure S5) and the Igraph network (Fig. 3B).

of nucleotide sequences significantly influence the accuracy of taxonomic assignments. BLASTx exhibited species-level accuracy of 65.2 per cent, 54.8 per cent, 49.4 per cent, and 41 per cent for sequence lengths of 1,000 nt, 500 nt, 200 nt, and 100 nt, respectively. For genus-level classifications of 1,000 nt, 500 nt, 200 nt, and 100 nt query sequences, the accuracies achieved were 99.4 per cent, 93.1 per cent, 90.6 per cent, and 86.7 per cent, respectively (Moubset et al. 2022). Taxonomic assignments were here carried out using DIAMOND searches, as mentioned above, at both the species and genus levels using sequence lengths of a minimum of 500 nt, 200 nt, or 100 nt with six distinct datasets being generated: S1, S2, and S5 for species level classifications of 100, 200, and 500 nt sequences and G1, G2, and G5 for genus level classifications of 100, 200, and 500 nt sequences (Supplementary Material).

Setting a threshold on the minimum number of reads needed for a virus to be considered as present

Read analyses of the negative and positive controls of each dataset (S1, S2, S5, G1, G2, and G5) revealed the degrees of contamination between samples that are likely to have arisen from either true environmental contamination or sequencing artifacts (e.g. Illumina-based index hopping, see below). To account for this baseline contamination, we considered that for each of the six datasets, any operational taxonomic unit (OTU; i.e. either species or genus) with an associated read count that was lower than the highest read count of identified contaminant OTUs was considered as not actually being present within the sample. For species datasets (S1, S2, and S5), thresholds for positive detection were

determined using read counts of viral species, while for genus datasets (G1, G2, and G5), thresholds for positive detection were determined based on the read counts of viral genera (Supplementary Material). In addition, any sequence with 100 per cent identity to the three viral isolates (sugarcane yellow leaf virus, sugarcane white streak virus, and sugarcane streak Egypt virus) infecting the positive controls of the American dataset as well as cacao swollen shoot Ghana Q virus (*Badnavirus epsiloninflatheobromae*) originating from an ongoing cocoa tree project of the research group were systematically removed from all samples in this study.

Lastly, we found virus sequences that were assigned to rice tungro baciliform virus (*Tungrovirus oryzae*) both in *B. barbinodis* and in the indigenous plant species sampled in France (*C. dactylon, B. phoenicoides*, and *D. glomerata*). This phytovirus has been recognized as an endogenous viral element (EVE): a viral sequence integrated into host genomes and transmitted from one generation to the next as if it was a normal cellular gene. Its presence in the rice genome suggests a possible broader integration within several species of the Poaceae family (Kunii et al. 2004; Vassilieff et al. 2023). All reads that were assigned to rice tungro baciliform virus sequence were therefore disregarded during further analyses. Similarly, reads assigned to two other potential EVEs assigned to the *Solendovirus* genus (Geering, Scharaschkin, and Teycheney 2010), tobacco vein clearing virus (*Solendovirus venanicotianae*) and sweet potato vein clearing virus (*Solendovirus venaipomeae*), were also discarded.

Accounting for Illumina-based index hopping

During library preparation, a unique index sequence was assigned to each sequencing library, which enabled pooling and sequencing of multiple libraries simultaneously during a single sequencing run. There is a probability that an index from one library can end up associated with reads from another library during demultiplexing reads. Consequently, those reads can be assigned to the incorrect library: a phenomenon commonly known as index hopping (Kircher, Sawyer, and Meyer 2012; Illumina 2017). Reads incorrectly attributed to a library due to index hopping are typically observed at significantly lower abundance compared to reads that are correctly assigned which can lead to incorrect interpretation of results during downstream analyses. To address this issue, the counts of reads and contigs with the same taxonomic assignment, sharing the same long multiplex identifier (MID) and sharing 100 per cent nucleotide identity, were recovered and compared between libraries. In libraries with the lowest abundance of these contigs/reads, the contigs/reads were considered as potentially resulting as a consequence of index hopping errors and were therefore disregarded in further downstream analyses (Supplementary Material).

Curation of NCBI nr protein database-based taxonomic assignments

As mentioned above, taxonomic assignments were achieved using the DIAMOND algorithm (Buchfink, Xie, and Huson 2015) implemented against the NCBI non-redundant (nr) protein database with an e-value threshold of <0.001. It is noteworthy that the NCBI non-redundant (nr) protein database uses the 'Taxonomy Browser Viruses - GenBank' database, which contained, at the time of analysis, 8,076 viral species encompassing 236,892 viral taxa, each of which had a unique 'Taxonomy ID'. Meanwhile, only 11,273 viral species are officially listed by the International Committee on Taxonomy of Viruses (ICTV) as per the latest April 2023 version of the Virus Metadata Resource (VMR MSL378 v1; https://ictv.global/vmr). Specifically, a total of 2,148 viral species are referenced as plant viruses in the VMR MSL378 list, including 84 species with secondary hosts such as invertebrates or fungi. In addition, the ICTV has mandated a significant shift in virus species nomenclature, requiring the adoption of a standardized binomial naming format ('Genus + freeform epithet'), leading to an unprecedented overhaul in virus species naming across the field (Walker et al. 2021). Consequently, several different viral species listed in GenBank may actually belong to the same viral species according to the ICTV (Supplementary Material). To address this issue, we chose the ICTV list as the gold-standard reference list of virus species and corrected through a five-step process the taxonomic assignments returned by the 'Taxonomy Browser Viruses' database (Supplementary Material):

1. For virus species with exactly matched species names in both ICTV and GenBank, no modifications were made, confirming the nomenclature accuracy of 6,837 viral species in GenBank, representing 40,427 distinct taxa within the GenBank database.
2. GenBank's delay in updating the names of viral species that have been modified in ICTV using the binomial system raised a significant challenge. A meticulous search was conducted to establish exact matches between GenBank viral species, which showed no correspondence in the first step, and the virus names from ICTV documented in the file 'VMR MSL378 v1' under the 'Virus name(s)' column. This correction step enabled the updating of 130 viral 'species', representing 132,033 distinct taxa within the GenBank database.
3. In GenBank, viral species are sometimes listed not under the 'Species' column but under the 'Tax_name' column, which represents viral strains. We conducted a search for exact matches between ICTV viral species and the taxa names in the 'Tax_name' column. This step enabled us to add 1,323 viral species, representing 8,160 distinct taxa within the GenBank database.
4. We also compared the former names of plant viruses from the ICTV under the 'Virus name(s)' column with the taxa names under the 'Tax_name' column of GenBank. This step allowed us to complement the list of viral species in GenBank with 1,786 additional viral species, representing 3,536 taxa within the GenBank database.
5. To further identify plant viral species, a final correction step was implemented by matching the accession numbers provided in the ICTV VMR with their corresponding accession numbers supplied in 'Taxonomy ID' in GenBank. This final correction step enabled the updating of 99 plant viral GenBank species names (out of the 219 unmatched GenBank plant viral species names).

Overall, our five steps of curation led to the identification of sequences in GenBank for 2,028 species of plant viruses out of the 2,148 currently established and listed in the ICTV VMR (94.41 per cent) (Supplementary Table S2). This list of 2,028 virus species names was used for scoring the presence of virus sequences. We obtained for each dataset (S1, S2, S5, G1, G2, and G5) a final incidence table inventorying the presence/absence of putative viruses identified in each plant sample. These six tables were used for all subsequent statistical quantitative analysis.

The list of 2,028 virus species names was also used to retrieve additional information (metadata) on the detected viral species.

Firstly, host range was retrieved and catalogued as follows. We identified several taxa classified as plant viruses according to the ICTV/GenBank which are, however, suspected to be mycoviruses, commonly found in fungi and categorized as plant viruses due to their isolation from plant samples. Indeed, several virus families contain both plant and fungal virus members, including *Metaviridae* and *Pseudoviridae* with reverse-transcribing genomes; *Chrysoviridae, Totiviridae*, *Reoviridae*, and *Partitiviridae* with dsRNA genomes; *Endornaviridae, Botourmiaviridae*, and *Mitoviridae* with (+)ssRNA genomes; and *Phenuiviridae* with (-)ssRNA genomes (Andika et al. 2023). While it is extremely difficult to determine whether a novel virus belonging to these families might have been infecting a fungus or a plant, we arbitrarily decided to consider viruses from these families to be mycoviruses. Note that, specifically in the case of alphapartitiviruses and alphaendornaviruses, phylogenetic analyses have shown that some lineages contain both plant and fungal viruses, suggesting that some plant viruses are more similar to fungal viruses than they are to other plant viruses, and vice versa (Andika et al. 2023). In addition, virus species assigned to the *Amalgaviridae* family were split into two groups: the members of the *Amalgavirus* genus that are only known to infect plant species were considered as plant viruses while the members of the *Zybavirus* genus that only infect fungi were considered as mycoviruses (Walker et al. 2019). Virus species were thus separated into two categories according to the host range of their corresponding viral genus: 'phytoviruses' (PHY) for confirmed plant viruses, and 'mycoviruses' (MYC) for those with potential fungal origins (Supplementary Table S3). We refer to these phytoviruses and mycoviruses collectively as plant-associated viruses (PAV). Finally, the putative mode(s) of transmission of each PAV identified in this study were inventoried according to current knowledge on the mode of transmission of members of each genus to which these PAV were taxonomically assigned (Supplementary Table S3). When a PAV was listed as being putatively both vertically and horizontally transmitted (Supplementary Table S3), it was counted twice (once as vertically transmitted and once as horizontally transmitted).

Prevalence and coinfection

General Linear Mixed Models (GLMM) were employed in R to compare viral prevalence and co-infection rates among plant species and countries, effectively compensating for imbalances in sampling between France and the USA (R Core Team, 2022). This approach accounted specifically for differences in the numbers of sites per continent, the number of indigenous plant species, and the number of samples per plant species. Viral prevalence, defined here as the proportion of samples of a given plant species containing at least one PAV at each site, showed overdispersion when binomial regression models were used (with function glmer() in R package lme4). It was therefore treated using a quasibinomial regression model with function glmmPQL() in the R package MASS. Co-infection, defined here as the number of distinct virus taxa (species or genus; OTU) infecting each plant sample, was first modeled using a Poisson regression with function glmer() in R package lme4. Overdispersion was detected for all tested models, and thus we applied quasi Poisson distributions with function glmmPQL() in the R package MASS. As statistical power was insufficient to test a single global model with all independent variables, we decomposed the problem into four complementary questions:

1. Do prevalence and coinfection of PAV OTUs (hereafter referred to as PAV) differ for *B. barbinodis* in its native and introduced ranges? The models included the country (i.e. France and Arizona) as a fixed effect and sampling site as a nested random effect, and were applied to the dataset containing only *B. barbinodis* samples (y ~ Country, random=~1|Site_name).
2. Do prevalence and coinfection of PAV differ between indigenous grass species in native and introduced ranges? The models included the country as a fixed effect and a random sampling site effect for each plant species and were applied on the dataset containing only indigenous plant samples (y ~ Country, random=~1|Site_name/Host_name). Note that *B. barbinodis* was included as an American indigenous plant species in this analysis. This comparison allowed the definition of a baseline viral pressure (i.e. the degree of PAV prevalence in plants at a given point in time and space) on indigenous grass species between the two countries.
3. Do prevalence and coinfection of PAV differ between *B. barbinodis* and indigenous plant species in its introduced range (France)? The models included the indigenous vs. exotic status as a fixed effect, and sampling site as a random effect and were applied to the dataset containing only France—derived samples (y ~ Bot_Ind, random=~1 | Site_name).
4. Do prevalence and coinfection of PAV differ between *B. barbinodis* and other indigenous plant species (i.e. excluding *B. barbinodis*) in their native range? The models included *B. barbinodis* vs. other indigenous plants as a fixed effect, and sampling site as a random effect and were applied to the dataset containing only American samples (y ~ Bot_Ind, random=~1 | Site_name).

Viral richness

Viral richness, defined as the number of distinct viral species or genera, was estimated and compared between continents and plant species using diversity analyses implemented in the package iNEXT (Chao et al. 2014; Hsieh, Ma, and Chao 2016). First, the viral richness of indigenous plants was compared between the native and introduced ranges (i.e. excluding *B. barbinodis* from France). To do so, incidence data from the 201 plant samples from USA and the 240 indigenous plant samples from France were pooled by country, resulting in a matrix of two columns corresponding to each country × 77 viral species containing the numbers of samples positive for each virus in each column. To cope with differences in sampling effort, viral richness was estimated (with confidence intervals) after rarefaction at 200 plants for both samples. Second, viral richness was compared between plant species in the native and introduced ranges of *B. barbinodis*. For this, we first excluded all plant species with less than 20 samples, leaving four plant species in the USA and five plant species in France, including *B. barbinodis* in both countries (hence a matrix of 9 × 73). Incidence data were then used to produce size-based rarefaction and extrapolation curves (along with confidence intervals of these) at 1, 10, 20, 30, 40, 50, 80, and 100 individuals per plant species. As there is no easy way to produce P-values for viral richness comparisons, viral richness was considered significantly different between two samples of a given sample size if the confidence intervals of the rarefaction and extrapolation curves did not overlap.

Plant—virus association analysis

To further dissect the complex interactions between plant species and their associated viral taxa in native and introduced ranges, we used a network approach. Venn diagrams provided insights into

unique and shared viral species among *B. barbinodis* and indigenous plant communities within each country. Sankey diagrams were implemented, using package 'networkD3' in R (Allaire et al. 2017), to explore the plant–virus associations at a species level. Each plant species was represented as a distinct node on the left side of the Sankey diagram. The width of each node was proportional to the number of individual plants sampled. From each of these 'plant species nodes', flow lines extended outwards connecting to various 'virus OTU nodes' based on observed interactions with the numbers of flow lines per plant species corresponding to the number of instances that a unique association between that plant species and a virus OTU was detected. Plant–virus networks were then constructed using igraph package in R (Csardi and Nepusz 2005) to highlight the distribution of viral taxa across host plant species and countries depending on viral characteristics such as host range and transmission mode.

Cloning and sequencing of the complete genomes of mastreviruses

To estimate the degree of OTUs overestimation caused by the incorrect assignment of contigs/singletons generated from unknown virus taxa, we decided to use the *Mastrevirus* genus (*Geminiviridae* family) as a model and we sequenced all mastrevirus genomes present in the plant samples towards this end. Hence, DNA was extracted from plant samples for which OTUs were assigned to the *Mastrevirus* genus using the DNeasy Plant Mini Kit (Qiagen, Germany). Extracted DNA was used as a template for PCR amplification of the complete genomes of mastreviruses using pairs of abutting primers designed based on assembled reads from the metagenomics step (Supplementary Table S4). PCR amplification and cloning of amplicons were performed as previously described (Candresse et al. 2014). The cloned genomes were Sanger-sequenced using primer walking (Azenta, Germany). Full genomes of the mastrevirus isolates thus determined are publicly available in GenBank (accession numbers: OR596401-OR596405).

Genetic and phylogenetic analyses

The five mastrevirus genome nucleotide sequences determined in this study and representative genomes of the *Mastrevirus* genus were aligned using MUSCLE (Edgar 2004). Pairwise identity analyses of complete mastrevirus genome nucleotide sequences were further carried out using SDT v1.2 (Muhire, Varsani, and Martin 2014). A maximum likelihood phylogenetic tree was constructed using PhyML3 (Guindon et al. 2010) implemented in MEGA11 with the HKY + G + I nucleotide substitution model chosen as the best-fit (Tamura, Stecher, and Kumar 2021). One hundred bootstrap replicates were used to test the support of branches.

Contigs from the USA and France grass samples that were assigned to the capsid gene of *Vicia cryptic virus* or the replication-associated protein gene of *Lolium latent virus* were translated to protein sequences and aligned using MUSCLE with representative genomes of the *Partitiviridae* and *Alphaflexiviridae* families, respectively. Contigs from the French grass samples that were assigned to the polyprotein gene of alphaendornaviruses were also translated to protein sequences and aligned with representative genomes of the *Alphaendornavirus* genus using MUSCLE. Neighbor joining phylogenetic trees were inferred from these protein alignments using MEGA11. One thousand bootstrap replicates were performed to quantify branch support.

Results and discussion

OTUs drawn from contigs overestimate actual plant-associated virus diversity

In this study, we aimed to compare viral OTU prevalence, richness, and composition of *B. barbinodis* and indigenous grass species from both the native area of *B. barbinodis* (Arizona, USA) and its introduced area (southern France) using a VANA-based metagenomics approach (Moubset et al. 2022). Trimmed Illumina-reads obtained from plant samples were *de novo* assembled into non-redundant contigs that were further taxonomically assigned through DIAMOND searches against the NCBI non-redundant (nr) protein database. In total, 4,987,513 cleaned reads from plants sampled in France (representing 7 per cent of the total number of 71,106,616 cleaned reads) were assembled into 44,047 contigs (contigs mean size = 260 nt, SD = 152 nt) that were further assigned to viral OTUs (Table 1). Similarly, 885,144 cleaned reads from plants sampled in the USA (representing 1 per cent of the total number of 82,327,473 cleaned reads) were assembled into 4,097 contigs (contigs mean size = 251 nt, SD = 693 nt) that were further assigned to viral OTUs (Table 1). On average, 222,208 (SD = 168,887) and 409,589 (SD = 440,321) cleaned viral reads were obtained per plant sample in France and in the USA, respectively, and 138 (SD = 1587) and 20 (SD = 34) viral contigs were recovered per plant sample in France and in the USA, respectively (Table 1). Several previous studies have pointed out that using all non-redundant contigs for taxonomic assignment tends to overestimate the total number of identified virus species, due to the incorrect assignment of contigs generated from unknown virus taxa (Aziz et al. 2015; García-López, Vázquez-Castellanos, and Moya 2015; Roux et al. 2017; McLeish et al. 2022). The present study is no exception, but in an attempt to address this issue, we aimed to estimate the degree of OTUs overestimation caused by the incorrect assignment of contigs/singletons generated from unknown virus taxa, using OTUs assigned to the *Mastrevirus* genus (*Geminiviridae* family) as a model. Fifteen contigs obtained from seven French individual plants from four plant species (*B. barbinodis, C. dactylon*, *B. phoenicoides*, and *D. glomerata*) collected from four sampling sites in France were assigned by DIAMOND searches to six mastreviral OTUs (*Bromus catharticus striate mosaic virus, Chloris striate mosaic virus, Digitaria ciliaris striate mosaic virus, Dragonfly-associated mastrevirus*, *Sorghum arundinaceum associated virus*, and *Switchgrass mosaic-associated virus*) when using dataset S2 (contigs >200 nt) (Supplementary Figure S2). Cloning and Sanger sequencing of the complete genomes of French mastreviruses yielded five apparently complete genome sequences (Supplementary Figure S3). Among these five genomes, two were determined from two plants of *B. barbinodis* and shared 78.6 per cent genome-wide identity against one another and should therefore be considered as belonging to the same species based on the ICTV *Mastrevirus* species demarcation threshold of 78 per cent genome-wide identity (Muhire et al. 2013) (Supplementary Figure S3). Both these *B. barbinodis*-derived mastrevirus genomes and two other sequences recovered from *B. phoenicoides* shared less than 78 per cent genome-wide identity with all other known mastreviruses and against one another, and thus they collectively represent three tentative new species within the genus *Mastrevirus*. In addition, one genome from *D. glomerata* shares 91.6 per cent genome-wide identity with *Sorghum arundinaceum associated virus*, and should therefore be classified as belonging to the *Mastrevirus* species *Sorghum arundinaceum associated virus*. Therefore, these analyses reveal that the six mastreviral OTUs would have more

accurately been classified into four *Mastrevirus* species (Supplementary Figure S2), suggesting that DIAMOND-based assignation of non-redundant contigs may yield a 50 per cent overestimation of the total number of *Mastrevirus* species. This also indicates that the taxonomic assignments based on DIAMOND searches are to be considered with caution and we therefore systematically added the suffix 'like' to the ends of inferred species-level classifications. Nevertheless, if we assume that the unknown viral OTU identification rate in wild plant species from unmanaged ecosystems is probably similar because these areas have been very poorly studied so far, we hypothesize that viral OTU overestimation should not be an issue for the purposes of the present study.

The richness and prevalence of plant-associated viruses in grasses from the Sonoran Desert (Arizona, USA) tend to be lower than in those from the Mediterranean (Southern France) setting

In this study, we first studied the metagenomics-based viromes of indigenous grasses growing in areas of the native and the introduced ranges of *B. barbinodis*. Importantly, the Arizona population of *B. barbinodis* used here shall neither be taken as representative of the whole native range (which is very broad), nor considered as the source population of the French invasion. Assuming that *B. barbinodis* was introduced to France via the wool trade, it is more plausible that its origin was Argentina (or neighboring countries of South America), which was an attested source of wool importations into Southern France over the 20th century (Angoulvent, 1948). Areas of the Sonoran Desert of Arizona (USA) were then used to illuminate how *B. barbinodis* compares to other grasses in a native setting. Besides, given the differences of ecological contexts between the Sonoran Desert of Arizona (USA) and the fallow lands colonized by *B. barbinodis* in southern France, determining the ecological processes that might impact the VRH was beyond the scope of the present study. Finally, all results presented below were highly consistent regardless of the dataset used (S1, S2, S5, G1, G2, and G5 datasets, Supplementary Table S5). Accordingly, dataset S2 (contigs >200 nt) was hereafter used for further analyses.

A total of 43 and 26 PAV were determined from the 240 French and 201 American indigenous plant samples, respectively, when using dataset S2 (contigs >200 nt) (Table 1). The GLMM further revealed that PAV prevalence in France (23.7 ± 5 per cent) was significantly higher (P-value = 0.0209) compared to PAV prevalence in the USA (9.9 ± 6 per cent) (Table 1; Supplementary Table S5). Note that prevalence values estimated by GLMM models, which took into account differences in sampling effort, are slightly different from observed prevalence values reported in Table 1. PAV prevalence of the French grasses in this study is comparable to previous estimates obtained from prairie ecosystems worldwide (Claverie et al. 2018). For example, in Oklahoma (USA) prairies, 25 per cent of plants carried identifiable plant viral sequences (Muthukumar et al. 2009) and in both the Western Cape region of South Africa and the Rhone delta region of France, 26 per cent to 36 per cent of plant samples harbored PAV sequences (Bernardo et al. 2018). To cope with differences in sampling effort, viral richness was compared between France and the USA after having removed all plant species with less than 20 samples and further having produced rarefaction and extrapolation curves (along with confidence intervals of these). Ecosystem-level PAV richness, rarefaction, and extrapolation curves were associated with large confidence intervals due to very low prevalence of each PAV in both ecosystems. Nevertheless, indigenous plant species in France exhibited both observed and extrapolated PAV richness that was higher than those of the indigenous plant species from the USA, albeit not significantly different (Table 1; Fig. 1A-B). For instance, examining *C. dactylon*, a plant species naturalized in both countries, highlights this contrast. In the USA, the viral richness for *C. dactylon* is quite low, with an extrapolated viral richness of three PAV for a population of 100 individuals. Conversely, in France, the same number of *C. dactylon* individuals exhibits a significantly higher estimation of viral richness up to 22 PAV, and no evident plateau in the curve (Fig. 1A, B). This significant disparity highlights how the viral pressure faced by the same plant species can vary markedly depending on the studied ecosystem. However, this result also emphasizes local-level variation within each country and the fact that unobserved interactions in all areas of the geographic range were not sampled. Viral richness was further estimated (with confidence intervals) after rarefaction at 200 plants for the 201 plant samples from the USA and the 240 indigenous plant samples from France. PAV richness of the French grasses was thus estimated at 38 ± 8 while the PAV richness of the USA plants was 24 ± 7 (Fig. 3A). Finally, PAV co-infection rates, i.e. estimates of the number of PAV identified from individual plants, were not significantly different between the two sampled ecosystems. PAV co-infection rates varied between one and four PAV for each infected indigenous plant from both France and the USA (Fig. 1C; Supplementary Table S5).

Overall, our study revealed that grasses from the French ecosystem were characterized by high PAV prevalence and richness, suggesting an increased viral infection pressure relative to the analyzed USA grasses. This result may reflect that the indigenous French grass populations were collected from fallow lands colonized by *B. barbinodis*. These ecologically disturbed areas are characterized by significant connectivity, richness, and abundance of plant species, which sharply contrasts with the Sonoran Desert setting in Arizona (USA) that over the past decade has experienced higher temperatures and less precipitation than historical averages and is characterized by scarce and disconnected indigenous plant populations.

Attenuation of virome release in invasive B. barbinodis populations approximately 50 years after their introduction in the Mediterranean basin

There were no significant differences in either viral OTU prevalence or coinfection rate in *B. barbinodis* between France and the USA, as could have been expected under the virome release hypothesis either for PAV or PHY (Fig. 1C-D; Supplementary Figure S4A-B; Supplementary Table S5). However, several results suggest that a virome release may have actually occurred. First, while viral OTU prevalence in *B. barbinodis* was significantly higher for both PAV (P-value = 0.0128) and PHY (P-value = 0.0232) than in other Poaceae species in the native area, PAV or PHY prevalence were similar in *B. barbinodis* as it was for indigenous species in the introduced area (Fig. 1D; Supplementary Figure S4A; Supplementary Table S5). In fact, PAV and PHY prevalence in *B. barbinodis* in the USA (20.7 ± 11 per cent and 17.2 ± 11 per cent) were higher than that of other indigenous species (5.6 ± 8 per cent and 4.1 ± 8 per cent), respectively. Conversely, PAV and PHY prevalence of *B. barbinodis* in France (15 ± 10 per cent and 3.7 ± 10 per cent) were lower than that of the four local indigenous species (23.7 ± 5 per cent and 10.4 ± 6 per cent), respectively, albeit not significantly different for both PAV (P-values = 0.1579) and PHY (P-values = 0.1360) (Table 1; Fig. 1D; Supplementary Figure S4A; Supplementary Table S5). Second, a similar pattern was observed for PAV and PHY richness. Indeed, PAV richness for *B. barbinodis* (whether estimated on rarefied samples or extrapolated counts) did not differ between

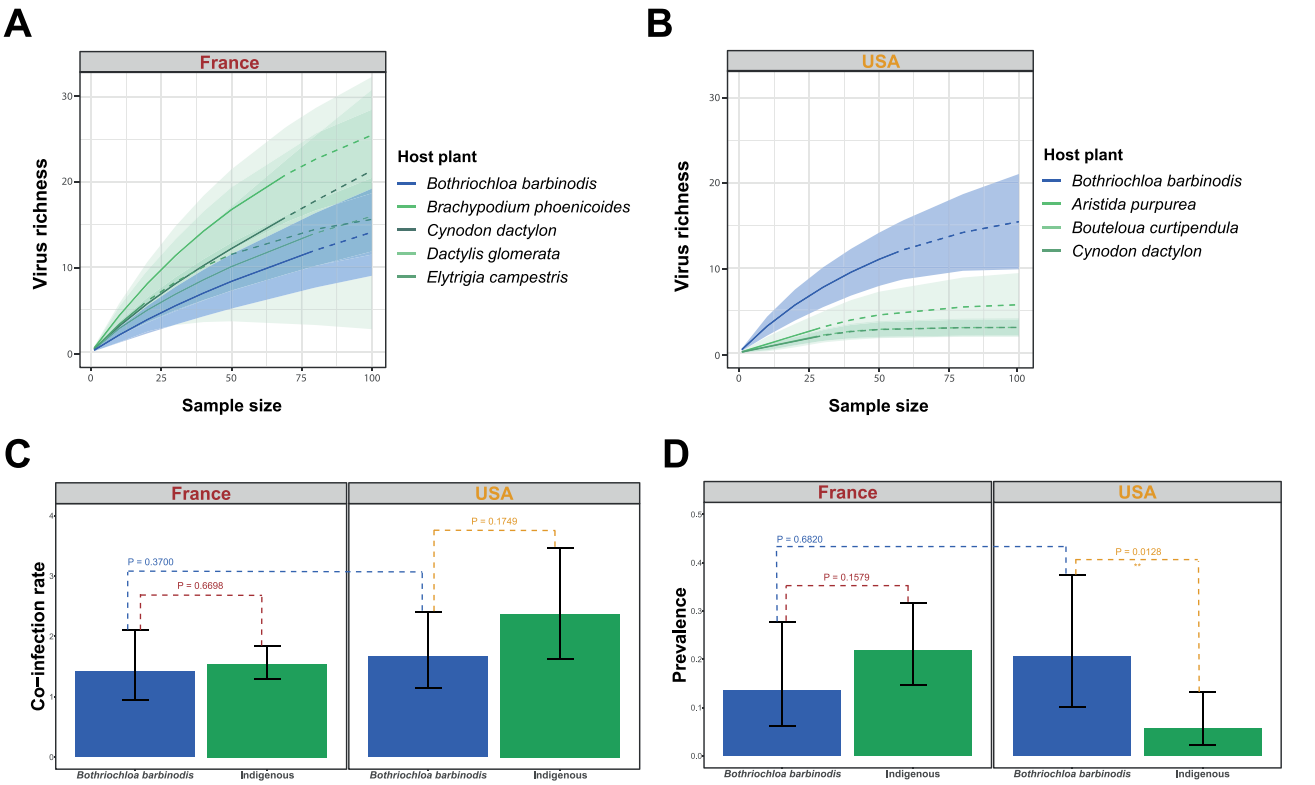

Figure 1. Comparison of PAV richness, prevalence, and co-infection rates between *B. barbinodis* and indigenous grass species in France and the USA. Rarefaction (solid lines) and extrapolation (dashed lines) curves, generated by the R package Inext (Chao et al. 2014; Hsieh, Ma, and Chao 2016), estimate PAV richness at 1, 10, 20, 30, 40, 50, 80, and 100 plants (along with confidence intervals) of each plant species where more than 20 individuals were collected in the field (A) in France and (B) in the USA. (C) Comparison of co-infection rates and (D) comparison of PAV prevalence between *B. barbinodis* and indigenous grass species in both the French and American ecosystems using generalized linear mixed models (GLMMs) with the R package lme4 (Bates et al. 2015).

the native and introduced areas (Fig. 1A, B, blue curves). However, while *B. barbinodis* had the highest associated PAV richness in its native area (Fig. 1B), its associated PAV richness fell below that of *D. glomerata*, which possessed the lowest PAV richness among the four sampled French indigenous species (Fig. 1A). Similarly, PHY richness, i.e. the part of PAV richness that is only counting plant viruses that overall are expected to be acute disease-causing viruses, for *B. barbinodis* and *C. dactylon* was estimated at 10 and 14 in France and 18 and 6 in the USA, respectively (Supplementary Figure S4C-D). These inverse trends in PHY richness for both plant species between the native and the introduced areas again suggest that *B. barbinodis* was less infected by plant viruses in its introduced area as compared to indigenous grass congeners, which is expected under the virome release hypothesis.

If the VRH follows the same trend as the ERH such that the IEPS has a temporary advantage that decreases as it colonizes the invaded environment and becomes a target of local virus species (Flory and Clay, 2013; Siemann, Rogers, and Dewalt 2006; Mitchell et al. 2010; Li et al. 2022), we posit that we may be witnessing a transitional stage out of a 'virome release' period for *B. barbinodis*, as this species was introduced to the French Mediterranean area between 42 and 54 years ago (Fried 2017). This interpretation is consistent with the results of a meta-analysis that demonstrated that herbivory and pathogen attacks on an IEPS intensify in the introduced range for about 50 years post-introduction of the species (Hawkes 2007).

Greater diversity and complexity in the French grass-virus network as compared to the USA

To further explore the broad clustering of viral communities by host species, we created a bipartite network (Dormann, Gruber, and Fründ 2008). These analyses revealed that the French network was denser and more complex (the main bipartite network descriptors are listed in Supplementary Table S6), overall displaying more interactions between PAV and plant species compared to the USA. In the French network, 12 out of 49 PAV (24.5 per cent) shared more than one host species, each infecting up to four different plant hosts. Conversely, in the American ecosystem, 3 out of 26 PAV (11.5 per cent) infected more than one plant host species (Fig. 2). These multi-host PAV were potentially mycoviruses in that they were all assigned to virus families—such as *Partitiviridae, Chrysoviridae*, *Genomoviridae*, and *Endornaviridae*—to which both PHY and MYC have been described. Consequently, we categorize these PAV as MYC, while other viral species, belonging to genera or families for which only plant hosts are known, were classified as PHY.

While PHY richness between the USA and the French ecosystems was not significantly different, MYC richness was significantly higher in the Mediterranean area relative to the Sonoran Desert setting (Fig. 3A). In fact, 40.8 per cent of PAV were categorized as MYC in France, as opposed to 26.9 per cent in the USA (Table 1; Supplementary Figure S5). Consequently, this overrepresentation of MYC in the Mediterranean basin may account

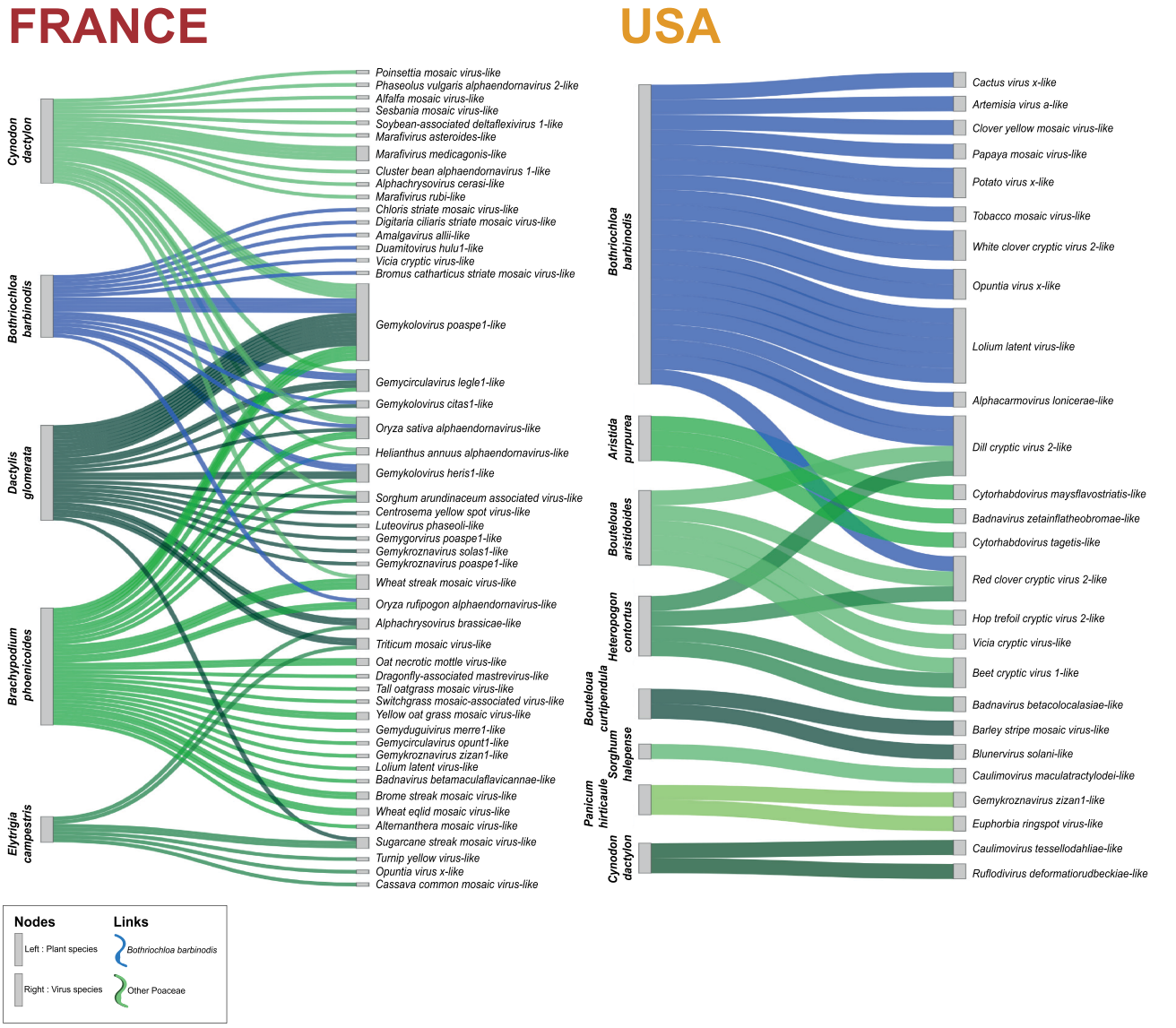

Figure 2. Sankey diagrams illustrating the plant–virus interactions (A) in France and (B) in the USA. The width of each node as well as the number of flow lines is proportional to the number of infected individuals of the same plant species.

for the complexity of the French plant-virus network, as well as for the disparity between both ecosystems observed in quantitative analyses evaluating viral richness (Supplementary Figures S4, S5, S6; Supplementary Table S5). In line with MYC richness, the observed prevalence of MYC was different between France (14.4 ± 5 per cent) and the USA (2.4 ± 6 per cent) while the observed prevalence of PHY were overall similar between both ecosystems, i.e. 8.8 ± 5 per cent in France and 7 ± 9 per cent in the USA (Table 1).

We undertook a complementary network approach which illustrated plant–virus interactions from a different perspective, allowing the identification of the nature of PAV infecting each sampled plant species in both ecosystems and their respective putative transmission modes. Representing the network of plant–virus associations in both the native and introduced areas enabled us to highlight the distinct viral compositions of both ecosystems, which had only four PAV in common—two PHY (*Opuntia virus X*-like and *Lolium latent virus*-like OTUs) and two MYC (*Vicia cryptic virus*-like and *Gemykroznavirus zizan1*-like OTUs) (Fig. 3B, Supplementary Figure S5). However, phylogenetic analyses revealed that the contigs assembled from the USA and French grass samples, both encoding the capsid protein of a virus with highest DIAMOND homology to *Vicia cryptic virus*, were actually not phylogenetically related. Indeed, the putative capsid protein sequence from the French grass grouped with that of *Vicia cryptic virus*, while the putative capsid protein sequence from the USA grass grouped very distantly with that of *Black grass cryptic virus 1*, another *Partitiviridae* family member (Supplementary Figure S7A). In addition, the contigs assembled from the USA and French grass samples that were both most closely related to the replication-associated protein of *Lolium latent virus* likely represent, in fact, two different new species in the *Alphaflexiviridae* family (Supplementary Figure S7B). Finally, no significant nucleotide similarities were found neither between the contigs assigned to *Opuntia virus X* nor between the contigs assigned to *Gemykroznavirus zizan1*, which have hampered any phylogenetic analyses of these sequences recovered from both the French and the American plant species.

Additionally, no statistical difference was found when comparing the viral OTU richness of putatively vertically and horizontally

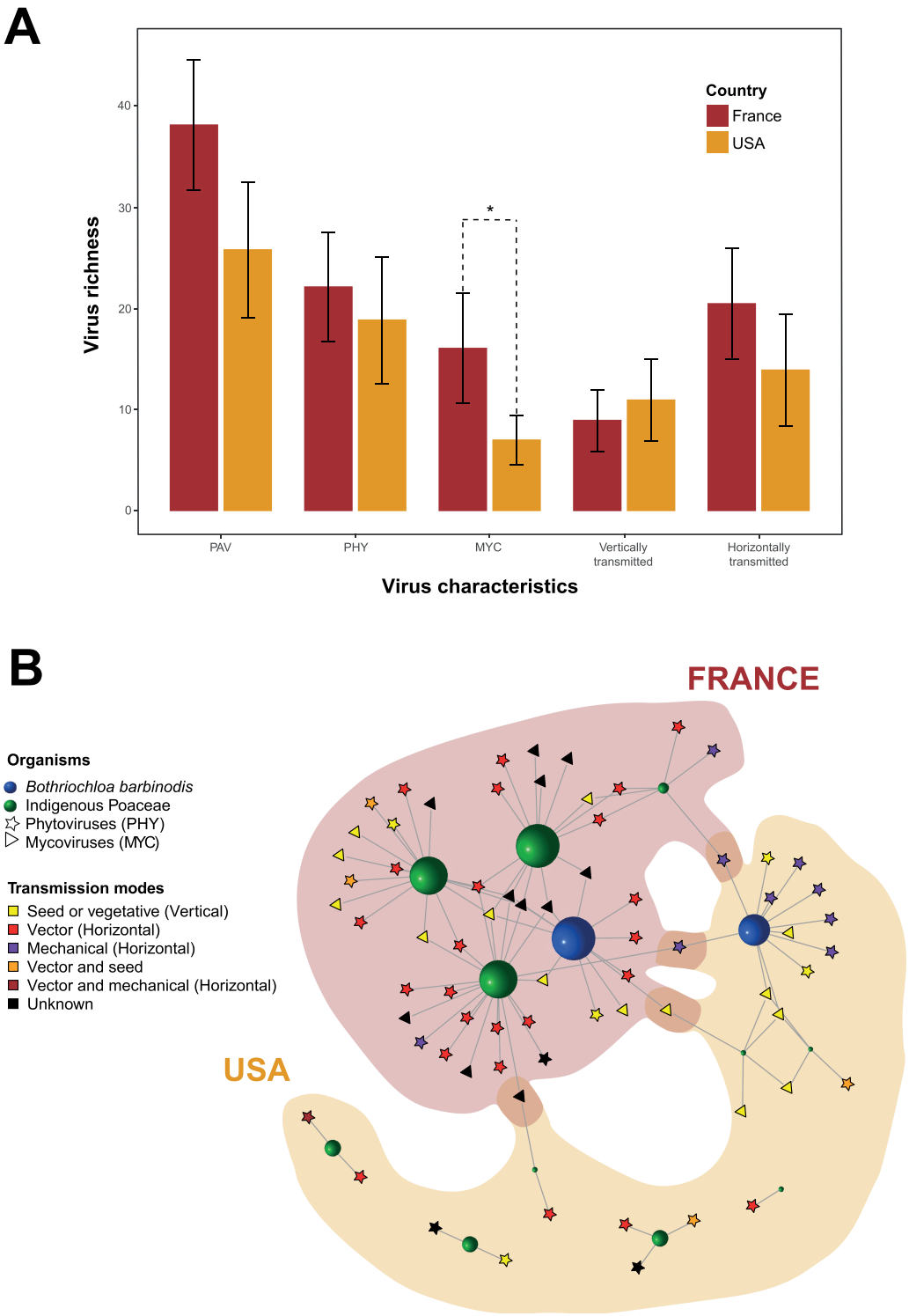

Figure 3. Comparative analysis of plant–virus interactions and viral richness in French *vs*. American ecosystems (A) Virus richness estimated (with confidence intervals) for PAV after rarefaction at 200 plants for the 201 plant samples from USA and the 240 indigenous plant samples from France. Virus richness is also estimated (with confidence intervals) according to the type of virus species (PHY or MYC) and according to the transmission modes (vertical and horizontal). Viral richness was considered significantly different between two samples of a given sample size if the confidence intervals of the rarefaction and extrapolation curves did not overlap (this is marked by an asterisk in the figure). (B) Network visualization of plant–virus interactions in French and American ecosystems: differentiation of PAV into PHY (plant viruses) and MYC (mycoviruses) with color-coded transmission modes and plant species demarcating indigenous species from *B. barbinodis*.

transmitted virus taxa between the French and the American ecosystems: In France, 13 virus species were putatively vertically transmitted, compared to 11 in the USA, while 26 virus species were putatively horizontally transmitted in France, as compared to 15 in the USA (Table 1; Fig. 3A-B). The remaining virus species, not included in these counts, have unknown transmission modes. However, it is noteworthy that in France the number of putatively horizontally transmitted virus species (26) was twice that of vertically transmitted virus species (13), a pattern not observed in the American ecosystem where the proportions were more similar (Table 1). The Igraph plot also indicated that PAV that are putatively vertically transmitted in the French ecosystem predominantly belonged to the group of MYC (Fig. 3B) that are overall known to either have host ranges restricted to a single species or host ranges restricted to a few closely related species (Pearson et al. 2009) even if some of them have a broader host range (Deng et al. 2022). The fact that several of these MYC were isolated from several grass species suggests, however, that they were horizontally transmitted. To test this hypothesis, we investigated the phylogenetic relationships of six contigs determined from *B. barbinodis, D. glomerata*, *B. phoenicoides*, and *C. dactylon*, encompassing the whole polyprotein of endornaviruses. Interestingly, these six contigs clustered together (Supplementary Figure S7C) and they most likely represent a new species in the *Alphaendornavirus* genus: a genus that is known to contain PHY that are transmitted solely via seed and pollen and MYC that can, in most cases, spread intracellularly, through hyphal anastomosis or sporulation (Hough et al. 2023). Hence, this result strongly suggests that this novel endornavirus is a mycovirus that was probably horizontally dispersed to the four grass species in France, most likely through fungal spores.

Reduced insect-borne and seed-borne transmission may have accounted for the virome release of invasive B. barbinodis in France

Eight putative MYC that spanned the *Partitiviridae, Mitoviridae, Endornaviridae*, and *Genomoviridae* families were identified from the French *B. barbinodis* population, six of which were found to be shared with other indigenous grass species, and one of which was shared with an American indigenous grass (Fig. 3B; Supplementary Figure S5). At the same time, four PHY from the *Amalgaviridae* and *Geminiviridae* families were uniquely isolated from the French *B. barbinodis* population (Fig. 3B; Supplementary Figures S5, S6). Surprisingly, besides one OTU assigned to the *Amalgaviridae* family, no other PHY with an RNA genome was identified in *B. barbinodis* from the French ecosystem even when they were surrounded by indigenous Poaceae species infected by viruses belonging to several genera of insect-vectored or mechanically transmitted RNA viruses, including viruses in the genera *Alfamovirus, Lolavirus, Luteovirus, Marafivirus, Poacevirus, Polerovirus, Potexvirus*, *Sobemovirus*, and *Tritimovirus*.

We already mentioned that the three OTUs assigned to the *Geminiviridae* family actually belong to a unique novel *Geminiviridae* species based on the ICTV *Mastrevirus* species demarcation threshold of 78 per cent genome-wide identity (Supplementary Figures S2 and S3). This novel mastrevirus is phylogenetically related to a group of mastreviruses infecting several Poaceae species in Australia (Supplementary Figure S3) (Kraberger et al. 2012). Since this group of mastreviruses has never been reported outside Australia, it is unclear how this novel mastrevirus ended up in southern France. Even if there is a strong suspicion that this novel mastrevirus was horizontally transmitted to *B. barbinodis* via an insect-vector, most likely a leafhopper (Fiallo-Olivé et al. 2021), the scenario of seed-transmission of this virus cannot be completely ruled out because at least one mastrevirus is suspected to be seed-transmissible (Qiao et al. 2020).

The second PHY identified in the French *B. barbinodis* plants shared 74 per cent amino acid identity with the partial RNA-dependent RNA polymerase (RdRp) of Allium cepa amalgavirus (*Amalgavirus allii*) that belongs to the *Amalgaviridae* family (Krupovic, Dolja, and Koonin 2015). Given that seed-transmission is the unique known mode of transmission of amalgaviruses, this virus identified in one *B. barbinodis* plant out of 80 is a candidate member of the original virome that may have been retained during the introduction and expansion of *B. barbinodis* in southern France. The effects of amalgaviruses and other vertically transmitted plant viruses (e.g. partitiviruses, endornaviruses, etc.) on their hosts are poorly understood but may be beneficial, including protection against acute infections of more harmful viruses, better tolerance to stress, regulation of nodulation, and increased plant height and fruit production (Takahashi et al. 2019).

In the context of rapid climate change, seed transmission may confer crucial benefits for plant viruses. For instance, it could allow virus survival under unfavorable conditions, such as prolonged drought periods (Pagán and Lozano-Durán 2022), a strategy that may be utilized by viruses in arid climates like those in Arizona, USA. However, only 8 of the 201 plants collected in the Sonoran Desert were infected by viruses related to viral species that are known to be seed-transmissible (including viruses in the *Alphachrysovirus, Alphapartitivirus, Betapartitivirus*, *Hordeivirus*, and *Tobamovirus* genera). None of these viruses was present in the French *B. barbinodis* plants. While alphachrysoviruses, hordeiviruses, and tobamoviruses are potentially acute disease-causing viruses, alphapartitiviruses and betapartitiviruses may potentially provide their hosts with survival benefits. These results suggest that vertically transmitted viruses rarely infect grasses in the Sonoran Desert, which may indicate that adaptation of these plants to the harsh environmental conditions of this setting has probably not been driven by the ubiquitous presence of beneficial vertically transmitted PAV.

However, it is important to note that these interpretations should be treated with caution, as we currently have only limited knowledge on the modes of transmission of plant viruses and the types of interactions they develop with their plant hosts. We therefore acknowledge that it is likely overly simplistic to hypothesize that vertically transmitted viruses are likely to be more beneficial (or less antagonistic) to their hosts than horizontally transmitted viruses. There is simply not enough available information to properly test such a hypothesis. Further, we acknowledge that there is also a high degree of uncertainty in the modes of transmission that we have assigned to each of the viral OTUs inventoried in this study in that these assignments are perhaps inappropriately inferred based on the assumption that the mode of transmission of an OTU will be the same as other better studied viruses belonging to the same genus. Here again, there is simply not enough available data on modes of transmission for the vast majority of known plant.

Nevertheless, our viral metagenomics-based study revealed that *B. barbinodis* was found associated with several mycoviruses in France where, contrary to Arizona, frequent and ubiquitous transmission of mycoviruses was observed. Interestingly, setting mycoviruses apart, only one putatively vertically transmitted phytovirus circulated among the 320 French *B. barbinodis* plants collected in this study. These results suggest that *B. barbinodis* in France did lose most of its native vertically transmitted PAV upon introduction. This would be congruent with (1) little or no seed

transmission of PAV and/or (2) a very low vertically transmitted PAV prevalence in the native population(s) as observed in Arizona. Similarly, only one putatively horizontally transmitted phytovirus infected these plants in the introduced area, pinpointing that *B. barbinodis* may have not yet gained new phytoviruses in the introduced ecosystem. This result implies a limited vector-mediated transmission of plant viruses to *B. barbinodis* in the introduced ecosystem, possibly as a result of a lack of insect vector preference for this newly introduced species (Liu and Stiling 2006; Meijer et al. 2016). Collectively, these results suggest that after experiencing an initial virome release phase following its introduction to France at least 42 years ago, *B. barbinodis* is probably transitioning, or has already transitioned, out of that phase and started interacting with local viruses. Whether this temporary release in the viral pressure over *B. barbinodis* populations was instrumental in its invasive success, e.g. through resource reallocation as stated in the original ERH, cannot be concluded from the present study, and will require identifying the source populations of the invasive French *B. barbinodis* and setting up proper life-history traits comparisons between native and invasive populations (Keane and Crawley 2002; Brian and Catford 2023).

Data availability

Illumina cleaned reads have been deposited in the sequence read archive of GenBank (accession number PRJNA721112 for the French plant samples and accession number PRJNA1020112 for the American plant samples). Full genomes of the mastrevirus genomes are publicly available in GenBank (accession numbers: OR596401-OR596405). Scripts and raw material were deposited in Zenodo with the following DOI: https://doi.org/10.5281/zenodo.10159099.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

The authors thank Frédéric Mahé for his excellent bioinformatics support.

Funding

O. Moubset and S. Ben Chéhida are recipient of Ph.D. fellowships from the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) and Agence Nationale de la Recherche (ANR) (Phytovirus project ANR-19-CE35-0008-02). P. Lefeuvre and J.-M. Lett were supported by the European Regional Development Fund (contract GURDT I2016-1731-0006632), the Conseil Régional de La Réunion, and CIRAD.

Conflict of interest: In the interests of transparency and to help readers to form their own judgments of potential bias, the authors declare no competing interests in relation to the work described.

References

Allaire, J. et al. (2017). *Package 'Networkd3,' D3 JavaScript Network Graphs from R.* <https://cran.r-project.org/package=networkD3> accessed 15 Sep 2023

Andika, I. B. et al. (2023) 'Cross-Kingdom Interactions between Plant and Fungal Viruses', *Annual Review of Virology*, 10: 119–38.

Angoulvent, P. (1948) 'Bilan de l'industrie lainiere en France Etudes Et Conjonctures - Union française/Economie Française 3ème Année N°6-8', 69–124.

Aziz, R. K. et al. (2015) 'Multidimensional Metrics for Estimating Phage Abundance, Distribution, Gene Density, and Sequence Coverage in Metagenomes', *Frontiers in Microbiology*, 6: 381.

Bankevich, A. et al. (2012) 'SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-cell Sequencing', *Journal of Computational Biology*, 19: 455–77.

Bartz, R., and Kowarik, I. (2019) 'Assessing the Environmental Impacts of Invasive Alien Plants: A Review of Assessment Approaches', *Neobiota*, 43: 69–99.

Bates, D. et al. (2015) 'Fitting Linear Mixed-Effects Models Using Lme4', *Journal of Statistical Software*, 67: 1–48.

Bernardo, P. et al. (2018) 'Geometagenomics Illuminates the Impact of Agriculture on the Distribution and Prevalence of Plant Viruses at the Ecosystem Scale', *The ISME Journal*, 12: 173–84.

Brian, J. I., and Catford, J. A. (2023) 'A Mechanistic Framework of Enemy Release', *Ecology Letters*, 26: 2147–66.

Buchfink, B., Xie, C., and Huson, D. H. (2015) 'Fast and Sensitive Protein Alignment Using DIAMOND', *Nature Methods*, 12: 59–60.

Candresse, T. et al. (2014) 'Appearances Can Be Deceptive: Revealing a Hidden Viral Infection with Deep Sequencing in a Plant Quarantine Context', *PLoS One*, 9: e102945.

Chao, A. et al. (2014) 'Rarefaction and Extrapolation with Hill Numbers: A Framework for Sampling and Estimation in Species Diversity Studies', *Ecological Monographs*, 84: 45–67.

Claverie, S. et al. (2018) 'From Spatial Metagenomics to Molecular Characterization of Plant Viruses: A Geminivirus Case Study', *Advances in Virus Research*, 101: 55–83.

Claverie, S. et al. (2023) 'Metagenomics Reveals the Structure of *Mastrevirus*–host Interaction Network within an Agro-ecosystem', *Virus Evolution*, 9: vead043.

Colautti, R. I. et al. (2004) 'Is Invasion Success Explained by the Enemy Release Hypothesis?', *Ecology Letters*, 7: 721–33.

Cottaz, C. et al. (2021) *Liste de Référence Des Plantes Exotiques Envahissantes de la Région Occitanie (V.1.0; 2021)* <https://doctech.cbnpmp.fr/pee-occitanie/liste-ref_plantes-exotiques-envahissantes_2021.pdf> accessed 15 Sep 2023.

Csardi, G., and Nepusz, T. (2005) 'The Igraph Software Package for Complex Network Research', *InterJournal Complex Systems*, 1695: 1–9.

Deng, Y. et al. (2022) 'Viral Cross-class Transmission Results in Disease of a Phytopathogenic Fungus', *The ISME Journal*, 16: 2763–74.

Dormann, C. F., Gruber, B., and Fründ, J. (2008) 'Introducing the Bipartite Package: Analysing Ecological Networks', *interaction*, 1: 8–11.

Edgar, R. C. (2004) 'MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput', *Nucleic Acids Research*, 32: 1792–7.

Elsheikh, E. A. E. et al. (2021) 'Role of Endophytes and Rhizosphere Microbes in Promoting the Invasion of Exotic Plants in Arid and Semi-Arid Areas: A Review', *Sustainability*, 13: 13081.

Faillace, C. A., Lorusso, N. S., and Duffy, S. (2017) 'Overlooking the Smallest Matter: Viruses Impact Biological Invasions', *Ecology Letters*, 20: 524–38.

Fiallo-Olivé, E. et al. (2021) 'ICTV Virus Taxonomy Profile: Geminiviridae 2021', *The Journal of General Virology*, 102: 001696.

Flory, S. L., Clay, K., and Thrall, P. (2013) 'Pathogen Accumulation and Long-term Dynamics of Plant Invasions', *Journal of Ecology*, 101: 607–13.

François, S. et al. (2018) 'Increase in Taxonomic Assignment Efficiency of Viral Reads in Metagenomic Studies', *Virus Research*, 244: 230–4.

Fried, G. (2017) 'À Propos de L'extension de *Bothriochloa Barbinodis* (Lag.) Herter En France', *Bulletin de la Société Botanique du Centre-Ouest*, 48: 119–26.

Gaertner, M. et al. (2009) 'Impacts of Alien Plant Invasions on Species Richness in Mediterranean-type Ecosystems: A Meta-analysis', *Progress in Physical Geography*, 33: 319–38.

García-López, R., Vázquez-Castellanos, J. F., and Moya, A. (2015) 'Fragmentation and Coverage Variation in Viral Metagenome Assemblies, and Their Effect in Diversity Calculations', *Frontiers in Bioengineering and Biotechnology*, 3: 141.

Geering, A. D., Scharaschkin, T., and Teycheney, P. Y. (2010) 'The Classification and Nomenclature of Endogenous Viruses of the Family Caulimoviridae', *Archives of Virology*, 155: 123–31.

Geisen, S. et al. (2021) 'Fungal Root Endophytes Influence Plants in a Species-specific Manner that Depends on Plant's Growth Stage', *Journal of Ecology*, 109: 1618–32.

Guindon, S. et al. (2010) 'New Algorithms and Methods to Estimate Maximum-likelihood Phylogenies: Assessing the Performance of PhyML 3.0', *Systematic Biology*, 59: 307–21.

Gutiérrez-Sánchez, Á. et al. (2023) 'Environmental Conditions Modulate Plant Virus Vertical Transmission and Survival of Infected Seeds', *Phytopathology®*, 113: 1773–87.

Hawkes, C. V. (2007) 'Are Invaders Moving Targets? The Generality and Persistence of Advantages in Size, Reproduction, and Enemy Release in Invasive Plant Species with Time since Introduction', *The American Naturalist*, 170: 832–43.

Hood-Nowotny, R. et al. (2023) 'Plant Invasion Causes Alterations in Darwin's Finch Feeding Patterns in Galápagos Cloud Forests', *Science of the Total Environment*, 895: 164990.

Hough, B. et al. (2023) 'Fungal Viruses Unveiled: A Comprehensive Review of Mycoviruses', *Viruses*, 15: 1202.

Hsieh, T. C., Ma, K. H., and Chao, A. (2016) 'iNEXT: An R Package for Rarefaction and Extrapolation of Species Diversity (Hill Numbers)', *Methods in Ecology and Evolution*, 7: 1451–6.

Hulme, P. E. (2009) 'Trade, Transport and Trouble: Managing Invasive Species Pathways in an Era of Globalization', *Journal of Applied Ecology*, 46: 10–8.

Illumina, I. (2017) *Effects of Index Misassignment on Multiplexing and Downstream Analysis*. <www.illumina.com> accessed 15 Sep 2023.

Jeong, S. et al. (2021) 'Invasive *Lactuca Serriola* Seeds Contain Endophytic Bacteria that Contribute to Drought Tolerance', *Scientific Reports*, 11: 13307.

Keane, R. M., and Crawley, M. J. (2002) 'Exotic Plant Invasions and the Enemy Release Hypothesis', *Trends in Ecology and Evolution*, 17: 164–70.

Kircher, M., Sawyer, S., and Meyer, M. (2012) 'Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform', *Nucleic Acids Research*, 40: e3.

Kraberger, S. et al. (2012) 'Australian Monocot-infecting Mastrevirus Diversity Rivals that in Africa', *Virus Research*, 169: 127–36.

Krupovic, M., Dolja, V. V., and Koonin, E. V. (2015) 'Plant Viruses of the Amalgaviridae Family Evolved via Recombination between Viruses with Double-stranded and Negative-strand RNA Genomes', *Biology Direct*, 10: 12.

Kunii, M. et al. (2004) 'Reconstruction of Putative DNA Virus from Endogenous Rice Tungro Bacilliform Virus-like Sequences in the Rice Genome: Implications for Integration and Evolution', *BMC Genomics*, 5: 80.

Lehan, N. E. et al. (2013) 'Accidental Introductions are an Important Source of Invasive Plants in the Continental United States', *American Journal of Botany*, 100: 1287–93.

Li, Y. X. et al. (2022) 'Diversity and Pathogenicity of *Alternaria* Species Associated with the Invasive Plant *Ageratina Adenophora* and Local Plants', *Peer Journal*, 10: e13012.

Liu, H., and Stiling, P. (2006) 'Testing the Enemy Release Hypothesis: A Review and Meta-analysis', *Biological Invasions*, 8: 1535–45.

Mack, R. N. et al. (2000) 'Biotic Invasions: Causes, Epidemiology, Global Consequences, and Control', *Ecological Applications*, 10: 689–710.

Maron, J., and Vilà, M. (2001) 'When Do Herbivores Affect Plant Invasion? Evidence for the Natural Enemies and Biotic Resistance Hypotheses', *Oikos*, 95: 361–73.

McLeish, M. J. et al. (2022) 'Metagenomics Show High Spatiotemporal Virus Diversity and Ecological Compartmentalisation: Virus Infections of Melon, Cucumis Melo, Crops, and Adjacent Wild Communities', *Virus Evolution*, 8: veac095.

Meijer, K. et al. (2016) 'A Review and Meta-analysis of the Enemy Release Hypothesis in Plant-herbivorous Insect Systems', *Peer Journal*, 4: e2778.

Mitchell, C. E. et al. (2010) 'Controls on Pathogen Species Richness in Plants' Introduced and Native Ranges: Roles of Residence Time, Range Size and Host Traits', *Ecology Letters*, 13: 1525–35.

Mitchell, C. E. and Power, A. G. (2003) 'Release of invasive plants from fungal and viral pathogens', *Nature* 421: 625–7.

Moubset, O. et al. (2022) 'Virion-associated Nucleic Acid-based Metagenomics: A Decade of Advances in Molecular Characterization of Plant Viruses', *Phytopathology®*.

Muhire, B. et al. (2013) 'A Genome-wide Pairwise-identity-based Proposal for the Classification of Viruses in the Genus *Mastrevirus* (Family *Geminiviridae*)', *Archives of Virology*, 158: 1411–24.

Muhire, B. M., Varsani, A., and Martin, D. P. (2014) 'SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation', *PLoS One*, 9: e108277.

Muthukumar, V. et al. (2009) 'Non-cultivated Plants of the Tallgrass Prairie Preserve of Northeastern Oklahoma Frequently Contain Virus-like Sequences in Particulate Fractions', *Virus Research*, 141: 169–73.

Pagán, I., and Lozano-Durán, R. (2022) 'Transmission through Seeds: The Unknown Life of Plant Viruses', *PLoS Pathogens*, 18: e1010707.

Pearson, M. N. et al. (2009) 'Mycoviruses of Filamentous Fungi and Their Relevance to Plant Pathology', *Molecular Plant Pathology*, 10: 115–28.

Pyšek, P. et al. (2020) 'Scientists' Warning on Invasive Alien Species', *Biological Reviews*, 95: 1511–34.

Pyšek, P., and Richardson, D. M. (2010) 'Invasive Species, Environmental Change and Management, and Health', *Annual Review of Environment and Resources*, 35: 25–55.

Qiao, Q. et al. (2020) 'Evidence for Seed Transmission of Sweet Potato Symptomless Virus 1 in Sweet Potato (Ipomoea Batatas)', *Journal of Plant Pathology*, 102: 299–303.

Rai, P. K., and Singh, J. S. (2020) 'Invasive Alien Plant Species: Their Impact on Environment, Ecosystem Services and Human Health', *Ecological Indicators*, 111: 106020.

R Core Team, R. (2022) 'R: A Language and Environment for Statistical Computing', R Foundation for Statistical Computing: Vienna, Austria.

Roossinck, M. J. (2011) 'The Big Unknown: Plant Virus Biodiversity', *Current Opinion in Virology*, 1: 63–7.

Roossinck, M. J. (2015b) 'Plants, Viruses and the Environment: Ecology and Mutualism', *Virology*, 479-480: 271–7.

Roossinck, M. J. et al. (2010) 'Ecogenomics: Using Massively Parallel Pyrosequencing to Understand Virus Ecology', *Molecular Ecology*, 19: 81–8.

Roossinck, M. J., and Schultz-Cherry, S. (2015a) 'Move Over, Bacteria! Viruses Make Their Mark as Mutualistic Microbial Symbionts', *Journal of Virology*, 89: 6532–5.

Roux, S. et al. (2017) 'Benchmarking Viromics: An in Silico Evaluation of Metagenome-enabled Estimates of Viral Community Composition and Diversity', *Peerj*, 5: e3817.

Rua, M. A. et al. (2011) 'The Role of Viruses in Biological Invasions: Friend or Foe?', *Current Opinion in Virology*, 1: 68–72.

Schönegger, D. et al. (2023) 'Carrot Populations in France and Spain Host a Complex Virome Rich in Previously Uncharacterized Viruses', *PLoS One*, 18: e0290108.

Seebens, H. et al. (2017) 'No Saturation in the Accumulation of Alien Species Worldwide', *Nature Communications*, 8: 14435.

Siemann, E., Rogers, W. E., and Dewalt, S. J. (2006) 'Rapid Adaptation of Insect Herbivores to an Invasive Plant', *Proceedings Biological Sciences*, 273: 2763–9.

Takahashi, H. et al. (2019) 'Virus Latency and the Impact on Plants', *Frontiers in Microbiology*, 10: 485724.

Tamura, K., Stecher, G., and Kumar, S. (2021) 'MEGA11: Molecular Evolutionary Genetics Analysis Version 11', *Molecular Biology and Evolution*, 38: 3022–7.

Vassilieff, H. et al. (2023) 'Endogenous Caulimovirids: Fossils, Zombies, and Living in Plant Genomes', *Biomolecules*, 13: 1069.

Vega, A. (2000) 'Revisión Taxonómica de Las Especies Americanas Del Género *Bothriochloa* (Poaceae: Panicoideae: Andropogoneae)', *Darwiniana*, 38: 127–86.

Vilà, M. et al. (2010) 'How Well Do We Understand the Impacts of Alien Species on Ecosystem Services? A pan-European, Cross-taxa Assessment', *Frontiers in Ecology and the Environment*, 8: 135–44.

Walker, P. J. et al. (2019) 'Changes to Virus Taxonomy and the International Code of Virus Classification and Nomenclature Ratified by the International Committee on Taxonomy of Viruses (2019)', *Archives of Virology*, 164: 2417–29.

Walker, P. J. et al. (2021) 'Changes to Virus Taxonomy and to the International Code of Virus Classification and Nomenclature Ratified by the International Committee on Taxonomy of Viruses (2021)', *Archives of Virology*, 166: 2633–48.

War, A. F. et al. (2023) 'Seed-endophytes Empower *Anthemis Cotula* to Expand in Invaded Range', *Current Plant Biology*, 34: 100281.

Virus Evolution, 2024, **10(1)**, 1–15
DOI: https://doi.org/10.1093/ve/veae025
Advance Access Publication 9 March 2024
Research Article