VIRUS EVOLUTION

Virus Evolution, 2024, **10(1)**, 1–16

DOI: https://doi.org/10.1093/ve/veae003
Advance Access Publication 5 January 2024
Research Article

New clades of viruses infecting the obligatory biotroph *Bremia lactucae* representing distinct evolutionary trajectory for viruses infecting oomycetes

Marco Forgia,[1,†] Stefania Daghino,[1,†] Marco Chiapello,[1,§] Marina Ciuffo,[1,**] and Massimo Turina[1,2,††,*]

[1]Institute for Sustainable Plant Protection, National Research Council of Italy, Strada Delle Cacce 73, Torino 10135, Italy and [2]Institute for Sustainable Plant Protection, National Research Council of Italy, Via Branze 39, Brescia 25123, Italy
[†]https://orcid.org/0000-0003-0101-1046
[‡]https://orcid.org/0000-0001-5722-1558
[§]https://orcid.org/0000-0001-7768-3047
[**]https://orcid.org/0000-0003-0422-408X
[††]https://orcid.org/0000-0002-9659-9470
*Corresponding author: E-mail: massimo.turina@ipsp.cnr.it

Abstract

Recent advances in high throughput sequencing (HTS) approaches allowed a broad exploration of viromes from different fungal hosts, unveiling a great diversity of mycoviruses with interesting evolutionary features. The word mycovirus historically applies also to viruses infecting oomycetes but most studies are on viruses infecting fungi, with less mycoviruses found and characterized in oomycetes, particularly in the obligatory biotrophs. We, here, describe the first virome associated to *Bremia lactucae*, the causal agent of lettuce downy mildew, which is an important biotrophic pathogen for lettuce production and a model system for the molecular aspects of the plant-oomycetes interactions. Among the identified viruses, we could detect (1) two new negative sense ssRNA viruses related to the yueviruses, (2) the first example of permuted RdRp in a virus infecting fungi/oomycetes, (3) a new group of bipartite dsRNA viruses showing evidence of recent bi-segmentation and concomitantly, a possible duplication event bringing a bipartite genome to tripartite, (4) a first representative of a clade of viruses with evidence of recombination between distantly related viruses, (5) a new open reading frame (ORF)an virus encoding for an RdRp with low homology to known RNA viruses, and (6) a new virus, belonging to riboviria but not conserved enough to provide a conclusive phylogenetic placement that shows evidence of a recombination event between a kitrinoviricota-like and a pisuviricota-like sequence. The results obtained show a great diversity of viruses and evolutionary mechanisms previously unreported for oomycetes-infecting viruses, supporting the existence of a large diversity of oomycetes-specific viral clades ancestral of many fungal and insect virus clades.

Keywords: mycovirus; oomycetes; bremia; lettuce; orphan proteins.

Importance

Viruses of the oomycetes are poorly studied, particularly in the case of obligate biotrophs for which the laboratory maintenance in axenic culture is impossible. This work shows for the first time a diverse set of viruses infecting the lettuce downy mildew *B. lactucae*. Additionally, it sets up a possible model organism to characterize different types of viruses that so far were only reported through metatranscriptomic studies, and for which no information of the host and clear phylogenetic and taxonomic position was known. We, here, describe unusual mechanisms of virus evolution that are at the root of the viral diversity associated to *B. lactucae*. The viruses here described are often specific to oomycetes, and are at the root of plant and animal virus clades, underling the possible role that Chromista played as origin of RNA virus biodiversity.

Introduction

Latest advances in virome investigations caused an outstanding increase in the number of viral sequences detected and characterized. In particular, homology-based approaches applied to public RNAseq data allowed the identification of around 10^5 novel viruses in two independent works (Edgar et al. 2022; Neri et al. 2022). Even if these studies often cannot allow a correct association between the virus and the host hampering any biological characterization, their great relevance relies on the phylogenetic implication for the evolutionary history of viruses and the availability of a large database of viral sequences that will be crucial for future virome studies.

In addition, it is still possible to unveil new groups of viruses that are too distant to be detected with homology-based approaches when the pipeline used to analyze the data is focused

on the open reading frame (ORF)an contigs (sequences encoding for putative proteins that don't have any hit when compared to a reference database). Examples of such findings are the splipalmiviruses and ambiviruses (Sutela, 2020; Forgia et al. 2021) and the ormycoviruses (Forgia et al. 2022), groups of mycoviruses that were undetectable through blast approach at the time of the first report and that demonstrates the limitation of homology-based searches to provide a complete picture of the RNA viral world. Despite this limitation, so far, the presence and monophyletic nature of the RNA-directed RNA polymerase (RdRp) still support it as a hallmark gene for RNA viruses (excluding the RNA viruses currently present in the realm Ribozyviria).

Many of the latest works demonstrate that fungi and oomycetes are often a great resource for the detection of new virus clades; the absence of an extracellular phase and the possibility to spread viruses among compatible vegetative groups often cause multiple viral coinfection in single fungal isolates (Marzano et al. 2016; Picarelli et al. 2019; Botella and Jung 2021) and several interesting viral taxa with new genomic features have been unveiled from these studies. Besides ormycoviruses and ambiviruses cited above, examples of these important discoveries are: (1) splipalmiviruses: the first mycoviral lineage where the viral replicase hosts the conserved palm-subdomains ABCDE on two different proteins (Sutela et al. 2020; Chiba et al. 2021), (2) the orfanplasmoviruses: bipartite viruses distantly related to narnaviruses (Chiapello et al. 2020), (3) the polymycoviruses (Kanhayuwa, 2015), and (4) the Yadokari viruses: replicase-encoding viruses that are replication- and encapsidation-dependent from an unrelated dsRNA virus (Das et al. 2021).

Mycoviruses are not only interesting for their diversity and for helping reconstruct the origin and evolution of viruses, but also for the study of their potential application in the control of plant pathogens. The most relevant example of this application is the hypovirus Cryphonectria hypovirus 1, causing hypovirulence in *Cryphonectria parasitica*, the fungal agent of chestnut blight disease. The presence of this mycovirus in the European population of *C. parasitica* allowed the control of the chestnut blight disease, as the infected fungus undergoes phenotypic changes that lead to a reduction in the ability to kill the plant (Turina and Rostagno 2007; Rigling and Prospero 2017).

After the discovery of virus-induced hypovirulence, many other cases of virus-induced effects were reported, showing that mycovirus could have hidden effects on the host grown in axenic culture on rich media that are instead observed when taking into consideration the relation between the virus, the host, and the environment. Some of these examples are the osmotic tolerance induced by mycoviruses from marine environment on a new fungal host (Nerva et al. 2017), the modification of the production of secondary metabolites induced by a mycovirus on *Aspergillus ochraceus* (Nerva et al. 2018), the influence of the mycovirus on the ratio between sexual and asexual reproduction in *Rhizopus microsporus* (Espino-Vazquez et al. 2020). Lastly, some studies showed that mycoviruses could cause a change in the behavior of different plant pathogens, inducing an endophytic lifestyle when infecting the normally pathogenic *Sclerotinia sclerotiorum* and *Pestalotiopsis theae* (Zhang et al. 2020; Zhou et al. 2021).

Most of the information regarding mycoviruses was obtained from studies on cultivable fungi and oomycetes because of their easy manipulation and maintenance; regarding oomycetes, the majority of the viruses detected were found in collections of different species related to the *Phytophthora* genus (Cai and Hillman 2013; Botella and Jung 2021; Poimala et al. 2021; Botella et al. 2022; Raco et al. 2022) and just few studies are focused on obligate biotroph-associated viruses: *Plasmopara viticola* (Chiapello et al. 2020), *Plasmopara halstedii* (Heller-Dohmen et al. 2011), and *Sclerophthora macrospora* (Yokoi et al. 1999; Yokoi, Yamashita, and Hibi 2003). Information for virus-induced phenotypes in oomycetes is really limited: only few effects on growth and sporulation are reported in literature (Poimala et al. 2022). With the aim of exploring the viruses infecting obligate biotrophs with important impact on agriculture, we investigated the virome of a collection of *Bremia lactucae*, which is a limiting factor for lettuce (*Lactuca sativa*) cultivation worldwide.

Bremia lactucae can penetrate lettuce cuticles and build haustoria from which the plant resources are exploited. Although it is an obligate biotroph, it is possible to maintain isolates in laboratory condition through inoculation of lettuce seedlings, making it a valuable model to study biotrophic lifestyles in *Oomycetes* (Michelmore and Wong 2008). Many effectors involved in the host–plant interactions are known for *B. lactucae,* and its control is obtained mainly through selection of resistant cultivars of lettuce, chemical compounds, and good agronomic practices (Parra et al. 2016); virus-caused hypovirulence was not yet reported in this species: a first step to develop a virus-based biocontrol tool requires an investigation of the virome associated to this species. Despite the abundance of information about this pathosystem, no viruses have been associated to *B. lactucae* to date. Furthermore, a precedent of virus-caused change in pathogen-resistance gene interaction was previously reported (Magnaporthe oryzae chrysovirus 1 strain A in differential races of *Magnaporte oryzae*): this prompted us to investigate a possible similar effect also in *B. lactucae* (Aihara et al. 2018). Aim of our study was the first characterization of the *B. lactucae*-associated virome and our results showed for the first time distinct mycoviral genome organizations that were not reported before; all the viruses described are the first reported for *B. lactucae*.

Results

Eleven isolates of *B. lactucae* from the Naktuinbouw collection were maintained on lettuce seedlings. Zoosporangia from infected seedlings were collected through water suspension and total RNA was extracted for each isolate studied. Total RNAs were then pooled in one sample that was analyzed through Illumina sequencing. The resulting library was cleaned as shown in previous works (Forgia et al. 2022) and assembled *de novo*. The metatranscriptome obtained was analyzed to retrieve RNA viruses through homology-based approaches and ORFan detection to identify putative viral contigs. This analysis allowed the identification of thirteen viruses and two putative RNA viral contigs encoding for ORFan proteins as they don't show conservation to any protein contained in the NCBI nr database. Identified viruses and contigs are shown in Table 1 together with information on the closest Basic Local Alignment Search Tool (BLAST) hit in the database and the number of reads mapping on the viral fragments. For each viral fragment, primers for quantitative reverse transcription-PCR (qRT-PCR) were designed and used for their detection in the eleven *B. lactucae* isolates. Based on qRT-PCR results (Table 2), each virus is associated to a specific distribution among the *B. lactucae* isolates; moreover, qRT-PCR analysis was conducted also on total nucleic acid extracted from the *B. lactucae* isolates to detect evidence of a possible genomic integration origin of the identified viral contigs or rare accumulation of DNA from an RNA virus replication cycle (Turina et al. 2018). No amplification of specific segments was obtained using total nucleic acid by PCR, implying that the nature of the identified

Table 1. List of viruses characterized in this work. For each virus, the total amount of reads mapping on the viral genome is shown together with the first blastp hit against the nr database (indicating query length and identity percentage of the result).

Virus	Acronym	Length	Total reads	Accession		BLASTp first hit	Identity	Query cover
Bremia lactucae associated endornavirus 1	BlaEV1	13,293	690,534	MN565677	ORF1	Phytophthora alphaendornavirus 1	37.2	93
Bremia lactucae associated dsRNA virus 1	BladsRNAV1	8,632	334,570	MN565678	ORF1	Plasmopara viticola lesion associated dsRNA virus 1	53.9	99
					ORF2	Plasmopara viticola lesion associated dsRNA virus 1	75	55.48
Bremia lactucae associated fusagravirus 1	BlaFV1	7,210	104,159	MN565680	ORF1	Phlebiopsis gigantea mycovirus dsRNA 2	31	65
					ORF2	Phytophthora castaneae RNA virus 4	28	64
Bremia lactucae associated yuevirus-like virus 1	BlaYlV1	6,427	53,279	MN565681	RNA1	Shahe yuevirus like virus 1	31.67	95
		1,625	9,442	MN565684	RNA2	Beihai sesarmid crab virus 3	24.81	58
Bremia lactucae associated ssRNA virus 1	BlaRNAV1	15,723	50,684	MN565682	ORF1	Xinjiang sediment flavi-like virus 1	26	44
Bremia lactucae associated partitivirus 1	BlaPV1	2,340	3,687	MN565687	RNA1	Plasmopara viticola lesion associated Partitivirus 7	89.03	100
		1,642	1,448	OR060922	RNA2	Plasmopara viticola lesion associated Partitivirus 7	86.83	98
Bremia lactucae associated ssRNA 2	BlaRNAV2	6,197	2,634	MN565688	ORF1	Crucivirus sp.	27.22	52
					ORF2	NA	NA	NA
					ORF3	NA	NA	NA
Bremia lactucae associated narnavirus 1	BlaNV1	2,953	183,263	MN565679	ORF1	Phytophthora infestans RNA virus 4	71.69	99
Bremia lactucae associated narnavirus 2	BlaNV2	3,315	28,614	MN565683	ORF1	Insect narna-like virus 2	40.37	73
Bremia lactucae associated splipalmivirus 1	BlaSPV1	2,560	1,923	MN565689	RNA1	Plasmopara viticola lesion associated narnavirus 2	98.24	100
		2,294	1,622	MZ926717	RNA2	Botrytis cinerea binarnavirus 3	57.45	94
		914	1,418	OR060921	RNA3	Sclerotinia sclerotiorum narnavirus 1	47.98	82
Bremia lactucae associated ditolivirus 1	BlaDTV1	5,038	962	MN565691	RNA1	Plasmopara viticola lesion associated totivirus-like 5	40	69
		4,053	1,818	MN565690	RNA2	Erysiphe necator associated totivirus 3	62	82
Bremia lactucae associated ORFan virus 1	BlaORFanV1	12,054	202,988	OR060923	NA	NA	NA	NA
Bremia lactucae associated yuevirus-like virus 2	BlaYlV2	4,038	9,368	MZ926716	RNA1	Plasmopara viticola lesion associated Yue-like virus 1	54.43	99
		3,142	8,304	OR060926	RNA2	NA	NA	NA
		1,876	23,272	OR060927	RNA3	NA	NA	NA
Bremia lactucae associated viral ORFan 2	BlaVORFan2	1,315	52,954	OR060924	ORF1	NA	NA	NA
Bremia lactucae associated viral ORFan 5	BlaVORFan5	1,550	11,540	OR060925	ORF1	NA	NA	NA

contigs is viral RNAs (data shown only for particularly interesting viral contigs in dedicated sections below). The identified viruses were further characterized for their genomic features and presented here based on phylogenetic/taxonomic criteria. Due to the limited space available, some of the viruses identified in this study are presented in Supplementary text: in specific, they are three viruses related to the narnavirids in the *Lenarviricota* (Bremia lactucae associated narnavirus 1, 2 and Bremia lactucae associated splipalmivirus 1), a partitivirus called Bremia lactucae associated partitivirus 1 and an endornavirus called Bremia lactucae associated endornavirus 1. Here below we will describe instead in more detail the viruses whose genomes present novel interesting features.

dsRNA viruses related to *Duplornaviricota*: first evidence of a new group of bipartite viruses

Three viruses that were detected through homology-based approaches appear to be related to *Duplornaviricota* phylum. Based on the closest viruses and on the phylogenetic analysis, we tentatively named the three viruses Bremia lactucae associated fusagravirus 1 (BlaFV1), Bremia lactucae associated dsRNA virus 1 (BladsRNAV1) and Bremia lactucae associated ditolivirus 1 (BlaDTV1) (Table 1).

BlaFV1 has a monopartite genome of 7,210 bp encoding for two putative proteins, where the putative protein on the 3′ end contains the RdRp palm-domain (thus called ORF1) and the 5′ ORF (called ORF2) encodes a putative protein of unknown function

Table 2. List of isolates of *Bremia lactucae* used in this work. For each isolate, the presence of any of the viruses discovered in this work is indicated with an x.

Virus name	NCBI	BL 17EU	BL 20EU	BL 21EU	BL 22EU	BL 23EU	BL 26EU	BL 31EU	BL 32EU	BL 33EU	BL 34 EU	BL 35EU
Bremia lactucae associated endornavirus 1	MN565677	x	x	x								
Bremia lactucae associated dsRNA virus 1	MN565678		x		x	x				x	x	x
Bremia lactucae associated fusagravirus 1	MN565680		x					x		x	x	x
Bremia lactucae associated yuevirus-like virus 1	MN565681	x						x			x	
	MN565684	x						x			x	
Bremia lactucae associated ssRNA virus 1	MN565682		x							x		
Bremia lactucae associated ssRNA virus 2	MN565688		x							x		
Bremia lactucae associated partiti virus 1	MN565687										x	x
	OR060922										x	x
Bremia lactucae associated narnavirus 1	MN565679		x	x				x				
Bremia lactucae associated narnavirus 2	MN565683		x	x								
Bremia lactucae associated splipalmivirus 1	MN565689								x			
	MZ926717								x			
	OR060921								x			
Bremia lactucae associated ditolivirus 1	MN565691						x					
	MN565690						x					
Bremia lactucae associated yuevirus-like virus 2	MZ926716									x		
	OR060926									x		
	OR060927									x		
Bremia lactucae associated ORFan virus 1	OR060923		x			x						
Bremia lactucae associated viral ORFan 2	OR060924		x	x				x				
Bremia lactucae associated viral ORFan 5	OR060925		x	x								

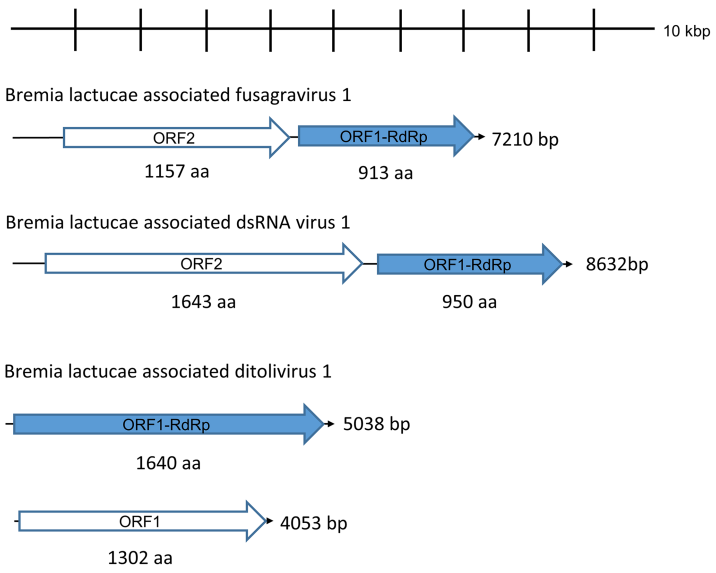


Figure 1. Schematic representation of the viruses related to the *Duplornaviricota* phylum. Black thin arrows represent the genomic segments and larger arrows represent the ORFs.

(Fig. 1). The closest blast hit to BlaFV1 putative proteins is shown in Table 1 together with identity values. The obtained identity value (below 32 per cent for each protein) is an evidence that BlaFV1 is a newly characterized virus and that it belongs to a new putative species related to the proposed group fusagravirus (Wang et al. 2016). This result is further confirmed by the phylogenetic analysis performed using sequences from different taxa related to *Duplornaviricota* phylum (Fig. 2) where BlaFV1 RdRp is found in a basal position of the clade containing the other fusagravirus-related viruses available in databases. The family 'Fusagraviridae' was informally proposed in 2016, and several fusagravirus-like viruses have been reported since then; since BlaFV1 shows low homology to its closest RdRp, we constructed an identity matrix using the RdRps from all the fusagraviruses included in the phylogenetic analysis. Results in Supplementary Fig. S1 show that, while a highly conserved group of fusagravirus can be detected (called 'group 1' in Supplementary Fig. S1), the conservation of the RdRp decreases heavily through the remaining viruses to reach less than 20 per cent for BlaFV1, by far the less conserved RdRp in this group, therefore supporting its basal position as outgroup to characterized fusagrivirids in the phylogenetic tree. This observation suggests that the fusagravirus we describe here could indeed be a member of a new clade of fusagravirid-like mycoviruses and not a bona-fidae fusagravirid.

BladsRNAV1 has an 8,623 bp long monopartite genome, encoding two ORFs called ORF1 (on the 3′ end, encoding a putative RdRp) and ORF2 (encoding a putative protein of unknown function) (Fig. 1). Both putative proteins show homology to the putative proteins from Plasmopara viticola lesion-associated dsRNA virus 1 (Table 1). Phylogenetic analysis places BladsRNAV1 in a still taxonomically unclassified group related to, but well distinct from, the genus *Botybirnavirus* (Fig. 2).

The most interesting results regarding dsRNA viruses are related to BlaDTV1. BlaDTV1 is a bipartite virus composed of an RNA1 fragment of 5,038 bp encoding for a putative protein of 1,640 amino acids showing a conserved RdRp palm-domain, and an RNA2 of 4,053 bp encoding for a putative protein of 1,302 amino acids with an unknown function (Fig. 1). The two RNAs were found together in the infected *B. lactucae* isolate BL26 (Table 2), and partial 3′ UTR end conservation could be reported even without carrying Rapid Amplification of cDNA Ends (RACE) analysis on the viral fragments (Supplementary Fig. S2) supporting the association of the two sequences being part of the same viral genome. Both putative proteins from BlaDTV1 RNA1 and RNA2 blasted against taxonomically unclassified toti-like viruses (Table 1); thus, the putative RdRp from BlaDTV1 was included in a phylogenetic analysis containing different viruses related to the Phylum *Duplornaviricota* (Fig. 2). The resulting phylogenetic tree (Fig. 2) grouped BlaDTV1 in a small clade containing two other viruses called Plasmopara viticola lesion associated totivirus-like 5 and Erysiphe necator associated totivirus 2 (PvlaTlV5 and EnaTV2), and these viruses are associated to biotrophic fungi and part of a bigger group of undefined toti-like viruses. PvlaTlV5 and EnaTV2 both show similar genomic organization compared to BlaDTV1 RNA1 (around 5 Kbp genome and a single ORF encoding for the putative RdRp), while the closer relative of this phylogenetic clade is called Hubei toti-like virus 5 (HTlV5), showing a bigger genome (around 7.5 Kbp) than BlaDTV1 RNA1 and a single ORF encoding for the putative RdRp. Interestingly, both proteins from BlaDTV1 RNA1 and 2 show homology when compared to HTlV5: as shown in Fig. 3, the putative protein from BlaDTV1 RNA2 has homology to the N-terminal of the one from HTlV5, while the putative protein from BlaDTV1 RNA1 has homology with the C-terminal (including the RdRp palm domain) of the HTlV5-encoded putative protein. This evidence suggests that BlaDTV1 is likely originated from the fragmentation of a larger monopartite viral ancestor related to HTlV5 and resulting in a bipartite genome; for this reason, we chose the name ditolivirus (divided toti-like virus) as a tentative name for BlaDTV1.

Since the phylogenetic analysis identified two viruses in the same group of BlaDTV1 (Fig. 2), we investigated the possible bipartite nature of PvlaTlV5 and EnaTV2 that were deposited as a single segment homologous to BlaDTV1-RNA1. Blastp analysis of the putative protein encoded by BlaDTV1 RNA2 showed that the first hits encoded putative viral proteins that were present in the same contig libraries where also PvlaTlV5 and EnaTV2 were found (Supplementary Table S1). We then tried to confirm an association between the other viruses grouped with BlaDTV1 (PvlaTlV5 and

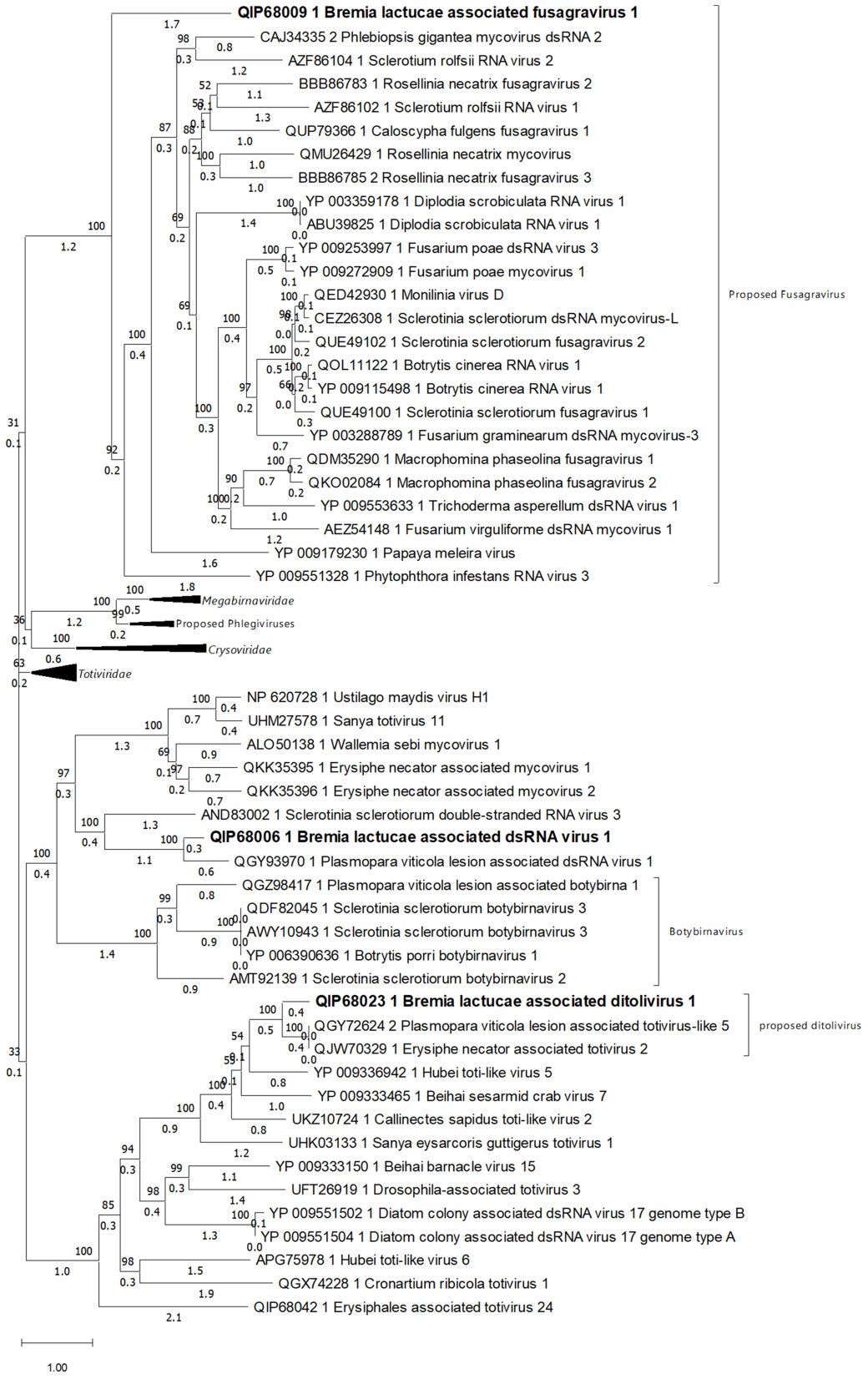
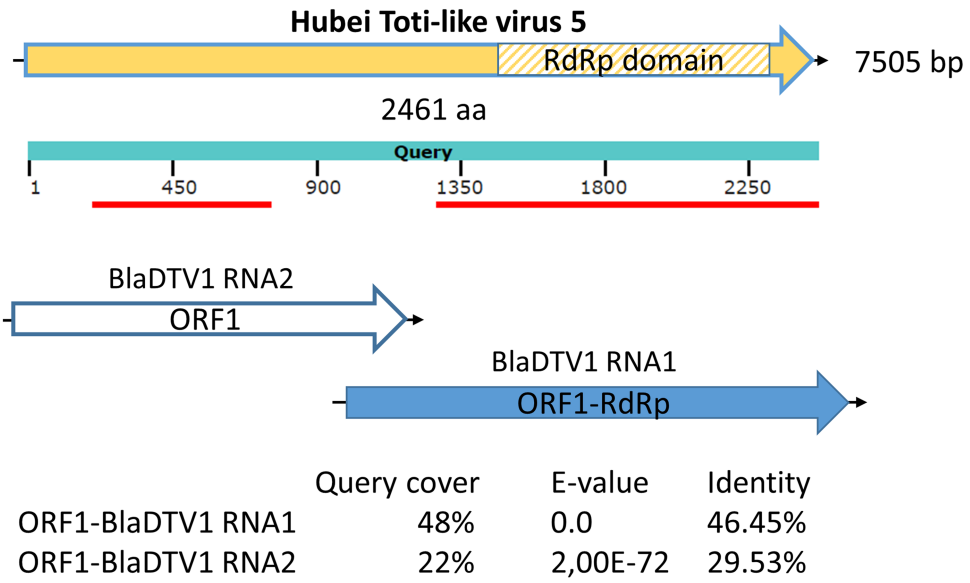


Figure 2. Phylogenetic analysis of the viruses related to the phylum *Duplornaviricota*. Members of the family *Megabirnaviridae*, *Crysoviridae*, *Totiviridae*, and the proposed group phlegiviruses are compressed. The viruses characterized in this work are displayed in bold. Sequences used for this tree are listed in Supplementary Table S6.



	Query cover	E-value	Identity
ORF1-BlaDTV1 RNA1	48%	0.0	46.45%
ORF1-BlaDTV1 RNA2	22%	2,00E-72	29.53%

Figure 3. Homology between the putative protein from Hubei Toti-like virus 5 and the putative proteins from Bremia lactucae-associated ditolivirus 1. The putative protein from HTlV5 was blasted against the putative proteins from the two genomic segments of BlaDTV1. The thin lines under the query box show the portion of the putative protein from HTlV5 where the homology with putative proteins from BlaDTV1 is found. Query cover, e-value, and identity percentage obtained for the two putative proteins from BlaDTV1 are displayed in the figure.

EnaTV2) and the viral fragments detected through blastp against BlaDTV1 RNA2 putative protein (called Plasmopara viticola lesion associated totivirus-like 3 and 4 and Erysiphe necator associated totivirus 3). For Erysiphe necator associated totivirus 3 (EnTV3), partial genomic end conservation could be reported for the 5′ end when compared to EnaTV2, supporting the bipartite nature of this virus (Supplementary Fig. S3A), but no data on the co-occurrence of the two fragments in the same biological sample are available since the work is not published yet. For Plasmopara viticola lesion associated totivirus-like 3 and 4 (PvlaTlV3 and 4), the co-occurrence of these fragments with PvlaTlV5 was reported in all the isolates found infected with PvlaTlV5, according to the data presented by Chiapello and coworkers (Chiapello et al. 2020). In this case, PvlaTlV3 and 4 both encode for BlaDTV1 RNA2-like proteins, but only one RdRp encoding fragment (PvlaTlV5) was found in the infected RNA samples. Since data on the study of the virome of *Plasmopara viticola* infected grapevines are available on the NCBI database, we conducted a focused search to detect possible other RdRp encoding segments that could support replication of this fragment, but only PvlaTlV5 could be found (data not shown). Nevertheless, partial conservation of the 5′ ends was reported between all the three fragments (PvlaTlV5, PvlaTlV3, and PvlaTlV4), suggesting that in this case the RdRp segment could be associated with two different versions of the putative RNA2 (whether PvlaTlV3 or PvlaTlV4) (Supplementary Fig. S3B). To clarify if the two RNA2 sequences (PlaTlV3 and 4) are always found associated to the RdRp encoding segment (PlaTlV5), we designed qRT-PCR primers for the three segments and we assayed single samples included in the library DMG-A from the work of Chiapello et al. Results in Supplementary Table S2 show that the three segments are always found together in infected single isolates, and we could never observe the presence of only one version of the RNA2 sequences coupled with the RdRp encoding segment. Nevertheless, we cannot exclude the remote possibility that the RdRp encoding segment PvlaTlV5 could be found in single host isolates coupled with only one form of RNA2 (PlaTlV3 or 4) since the RNA samples analyzed in this work were obtained from downy mildew lesions on grapevine leaves, thus possibly (but unlikely given the lesion infection process) resulting from more than one oomycetous isolate. Taken together, these results provide the first evidence of a new group of bipartite dsRNA viruses likely resulting from a splitting event of a larger virus. One member of this group (PvlaTlV5) also shows evidence of different variants of the RNA2 fragment maintained in the same isolate and therefore evidence of a further shift from bipartite to tripartite.

The last interesting evidence that we could observe about the newly identified ditoliviruses was the distribution of reads mapping on the positive and negative sense of the viral genome. Results displayed in Supplementary Table S3 show that, for the ditoliviruses where HTS data were available (BlaDTV1 and PvlaTlV5), a larger number of reads mapping on the negative sense of the viral genomes is present (with a variable ratio in different libraries, but always above 4), with RNA2s always showing a higher number of total reads mapping on them (around twice the amount of reads mapping on the RNA1). This evidence, confirmed for two different viruses with similar features, supports the hypothesis that the whole group of ditoliviruses could exploit replication mechanisms that rely on the accumulation of more negative sense genome strand as was already described for other dsRNA viruses (Dave et al. 2019; Chiapello et al. 2020).

Yue-like negative ssRNA viruses

Blast searches allowed the identification of two fragments encoding for putative proteins with homology to RdRp from negative sense RNA viruses called Yuevirus, identified by Shi and coworkers in invertebrate metasamples (Shi et al. 2016). This relatively new group of viruses hosts a small number of putative species and is phylogenetically related to a different taxon called Qinvirus. Moreover, a recent work on the investigation of the virome of *Plasmopara viticola*-associated lesions on grapevine identified a group of fungal/oomycetes Yue-associated viruses (Chiapello et al. 2020). The two viruses identified by our work are called Bremia lactucae associated yuevirus-like virus 1 and 2 (BlaYlV1, BlaYlV2).

The phylogenetic analysis performed on the putative RdRp suggests a close relation between bona fide Yueviruses and BlaYlV1 (Fig. 4A). A second fragment associated to BlaYlV1 was also detected through homology-based approaches, showing conservation with Yuevirus putative protein encoded by RNA2 (Fig. 4B). BlaYlV1 is therefore proposed as a new Yuevirus with a bipartite genome. BlaYlV2 instead is found in the *P. viticola*-related group (Fig. 4A) and no additional fragments could be found through homology-based methods. However, the ORFan pipeline detected two fragments that showed partial conservation of the 5′ and 3′ ends compared with BlaYlV2 (Fig. 4B, Supplementary Fig. S4). qRT-PCR reactions confirmed the presence of these two fragments in the *B. lactucae* isolate infected with BlaYlV2, further confirming their nature as RNA2 and RNA3 of this new Yue-like virus (Table 2) with a tripartite genome where RNA2 and RNA3 host putative proteins of unknown function showing no homology to any sequence in the NCBI nr database. It is particularly relevant that the Yue-like viruses in the group hosting BlaYlV2 and the Plasmopara viticola lesion-associated Yue-like viruses were initially reported with monopartite genomes only encoding for putative RdRp as no conservation is detected on RNA2 and RNA3 putative protein. Thus, a focused search was performed using putative proteins from BlaYlV2 RNA2 and RNA3 to detect putative additional fragments for the viruses clustering with BlaYlV2 (Chiapello et al. 2020). With this analysis, two fragments encoding for putative proteins sharing homology to BlaYlV2 RNA2 and RNA3 were identified. The RNA2 homolog is a 3,310 bp long contig; the putative protein encoded has an identity of 41.27 per cent (evalue 0.0, query cover 97 per cent) with the one encoded by BlaYlV2 RNA2. The RNA3 homolog is a 1,527 bp contig encoding for an incomplete putative protein with an identity of 42.11 per cent (evalue 0.0, query cover 80 per cent) with the putative protein encoded by BlaYlV2 RNA3 (Fig. 4B). Comparison of the 5′ and 3′ ends of the putative RNA2 and RNA3 identified in the *Plasmopara*-related libraries showed conservation of the 3′ with the PlaYlV1 sequence and partial conservation of the 5′, as the putative RNA3 from PvlYlV1 is incomplete, supporting an association among the three contigs representing the putative complete genome of this virus belonging to a new clade of tripartite Yue-like viruses including BlaYlV2 (Supplementary Fig. S4).

On the contrary, it was impossible to detect any fragment associated with Plasmopara viticola lesion-associated Yue-like virus 2 and 3 through comparison using RNA2 and RNA3 putative protein from BlaYlV2. This, together with phylogenetic analysis placing PvlaYlV2 and 3 relatively distant from BlaYlV2 and PvlaYlV1, suggests that these viruses would probably belong to at least two distinct taxonomic groups, and information about possible fragments associated to the RdRp-encoding segment is still lacking.

Positive ssRNA-related viruses unveil the first permuted RdRp in mycoviruses

Through blast similarity search approaches, we identified two viruses related to positive sense ssRNA viruses that we named Bremia lactucae-associated ssRNA virus 1 and 2 (BlaRNAV1 and BlaRNAV2).

BlaRNAV1 is 15,723 bp long, and it encodes a putative protein of 4,709 amino acids showing a conserved RdRp domain and a helicase domain (Fig. 5A) with similarity with flavi-like viruses detected in metatranscriptomic works and with a virus called Diatom colony associated ssRNA virus 1 (Table 1) (Urayama, Takaki, and Nunoura 2016). Interestingly, the helicase domain from BlaRNAV1 putative protein has its closest hit (besides the unidentified viruses mentioned before) among the potyviruses (phylum *Pisuviricota*), whereas the RdRp domain shares the highest amino acid identity with the pestiviruses (members of the family *Flaviviridae,* phylum *Kitrinoviricota*). Taking into account the relatively low number of viruses blasting against BlaRNAV1 and the low homology shown by the blast results (the closest hit has an identity of 26.63 per cent with only 44 per cent query cover according to the blastp analysis showed in Table 1), phylogenetic analysis was not performed for this virus that likely belongs to a new undefined group of viruses that could be originated from a rearrangement event between phylogenetically distant viruses. To confirm whether the *in silico* assembly of BlaRNAV1 was correct, we performed overlapping PCR on almost the complete viral genome: both the presence of PCR fragments of the expected size and their sanger sequencing of each segment confirmed the data obtained from the bioinformatics approach (Supplementary Fig. S5). Finally, confirmation of the absence of BlaRNAV1 in the DNA fraction of the infected *B. lactucae* isolate was obtained through comparison of the RT-PCR on the cDNA and on the total nucleic acid extract from BL20EU isolate (Supplementary Fig. S6).

BlaRNAV2 is 6,197 nt long, encoding three ORFs called ORF1, ORF2, and ORF3 (Fig. 5B). ORF1 encodes a putative protein with homology to tombusvirus coat proteins, while ORF2 and ORF3 showed no homology to any protein in the database. The presence of the contig was confirmed through qRT-PCR in single isolates of *B. lactucae* (Table 2) and no evidence of DNA associated to BlaRNAV2 was detected, thus confirming the putative viral nature of the contig (Supplementary Fig. S6). Overlapping PCR on nearly the full BlaRNAV2 sequence assembled by the HTS analysis (Supplementary Fig. S7) confirmed the correspondence to *in silico* assembly. Since no clue on possible association between BlaRNAV2 and other mycoviruses detected in the infected isolates could be hypothesized, and no hit could be found through blastp analysis for both ORF2 and ORF3, we performed protein structure homology-modeling using Swiss-model tool for ORF2 and ORF3. Results gave no relevant hit for ORF2, but ORF3 showed a hit against the alphapermutotetravirus Thosea asigna virus RdRp (with a QMEANDisCo value of 0.38 ± 0.06). Permutotetraviruses are characterized by an RdRp protein that conserves the typical folded structure of an RdRp while the primary structure displays a swap between the A and C motifs of the conserved palm domain. Further evidence of the permuted order of the conserved domains from the RdRp was obtained through alignment between BlaRNAV2 ORF3 sequence and a number of permuted RdRp sequences retrieved from NCBI (Fig. 5C). As shown by the picture, the conserved A, B, and C motifs are found in the same order between BlaRNAV2 ORF3 and other permuted RdRps, supporting the evidence of the first identification of a permuted RdRp in Chromista-associated viruses. Even if the permuted palm domain can be observed through sequence alignment between BlaRNAV2 RdRp sequence and other permuted viruses, the homology between these proteins is too low to be detected through blastp approach, suggesting an independent origin of the oomycetes-related permuted mycovirus. Even if we could not get hits when blasting BlaRNAV2 RdRp against the NCBI nr database, some hits could be obtained through tblastn against the Riboviria database (Neri et al. 2022). In this case, the majority of the RdRp blasting with BlaRNAV2 were characterized as belonging to the Phylum *Kitrinoviricota*, with only few exceptions that were taxonomically uncharacterized (data not shown), thus suggesting a possible phylogenetic origin for BlaRNAV2 permuted RdRp linked to this clade of positive sense ssRNA viruses.

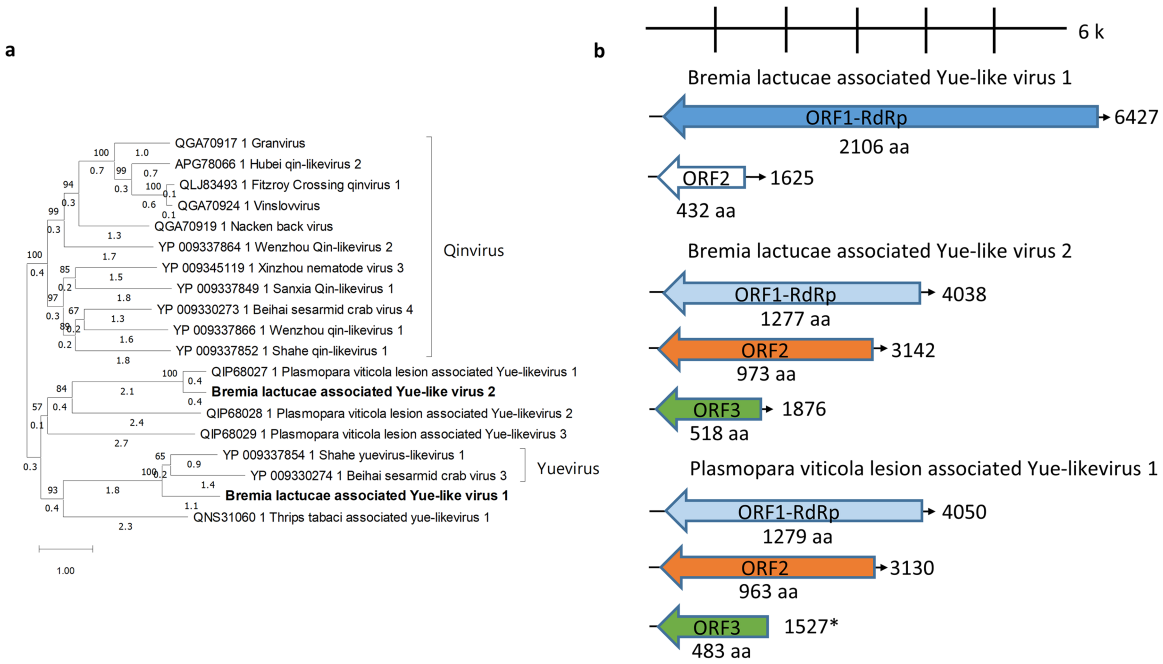


Figure 4. Phylogenetic analysis and genomic representation of yue-like viruses. Panel (A) shows the phylogenetic tree obtained using members of the yue-like and qin-like viruses available in the NCBI database (Supplementary Table S6). The viruses characterized in this work are represented in bold. Panel (B) shows the schematic representation of the Bremia lactucae-associated Yue-like virus 1 and 2 and the putative complete genome of Plasmopara viticola lesion-associated Yue-like virus 1 that was found through homology comparison with BlaYlV2. The RNA2 and RNA3 of Plasmopara viticola lesion-associated Yue-like virus 1 were deposited in NCBI under the accession numbers OR387775 and OR387776, respectively. Thin arrows represent the genomic segments and larger arrows represent the putative proteins.

Because of the novelty of BlaRNAV1 and BlaRNAV2, we decided to further demonstrate the specific association with *B. lactucae* as a natural host. Using the *B. lactucae* isolate BL20EU (infected with both BlaRNAV1 and BlaRNAV2), we inoculated lettuce seedlings and collected samples of (1) zoosporangia, (2) non-inoculated lettuce seedling, (3) inoculated seedling without zoosporangia clearly emerging from the cotyledons, and (4) inoculated seedling where BL20EU zoosporangia were already present on its surface. qRT-PCR were performed on the RNA extracted from these samples to show that the two viruses accumulate in the zoosporangia sample and are not detectable in the non-infected lettuce seedlings (Fig. 6), thus supporting that BlaRNAV1 and 2 are viruses infecting *B. lactucae*.

Viral ORFan contigs

ORFan investigation was performed to identify putative viral contigs that are too distant to be detected with canonical homology-based approaches. In our HTS library, five orfan contigs were selected and qRT-PCR was performed on cDNA and DNAse-untreated total RNA to demonstrate their nature as RNA species-only, thus confirming their putative viral nature (Supplementary Fig. S6). Two ORFan contigs were associated to BlaYlV2 and have been described above. Two sequences named Bremia lactucae-associated viral ORFan 2 and Bremia lactucae-associated viral ORFan 5 (BlaVORFan2, BlaVORFan5) were confirmed as putative viral RNA sequences since no DNA associated was detected through q-PCR. Moreover, even if the presence of BlaVORFan2 and 5 is the same observed for the narnaviruses Bremia lactucae narna like virus 1 and 2, respectively (Table 2), these two ORFans could not be clearly associated to any of the viruses detected in this study as no evidence of conservation on the ends could be provided. BlaVORFan2 is a 1,315 nt long sequence encoding for a putative protein of 197 aa while BlaVORFan5 is a 1,550 nt long sequence encoding for a putative protein of 282 aa (Fig. 7A). We derived the predicted structural features of both ORFs-encoded proteins with Alphafold2 (Supplementary Fig. 8A). We then interrogated structural protein databases to look for conserved features with both DALI and Foldseek search server. An interesting hit for BlaVORFan2 protein was the carboxy terminal domain present in BlaVORFan2 and two ORFan proteins present in *Bremia lactucae* (CCR75_002592) and *Peronospera halstedii* (PHALS_00922). The two proteins are conserved between themselves but no other hits are present when searching the databases using them as queries; using matchmaker in the Chimera X software, we could confirm the putative structural conservation of this carboxy terminal domain (Supplementary Fig. 8B). Abundance distribution was investigated for BlaVORFan2 using the same approach proposed for BlaRNAV1 and 2, and the evidence of a correlation between the infected oomycetes and BlaVORFan2 was observed (Supplementary Fig. 8C). The presence of the gene locus encoding for the *B. lactucae* CCR75_002592 protein was checked in the infected isolate BL20EU (used for the distribution experiments and infected also with BlaVORFan2), confirming its detection both in cDNA and total nucleic acid and thus demonstrating the genomic nature of CCR75_002592 (Supplementary Fig. S6), contrary to what happens with BlaVORFan2. To further exclude the association between BlaVORFan2 and BlaNV1, RACE analysis were performed to obtain the complete sequence of BlaVORFan2 excluding conservation on the 5′ and 3′ ends compared to BlaNV1.

The most interesting sequence identified through the ORFan pipeline was called Bremia lactucae associated ORFan virus 1 (BlaORFanV1). This sequence is 12,054 nt long, encoding for four putative proteins called ORFA (1702 aa), ORFB (1476 aa), ORFC (161 aa), and ORFD (175 aa) (Fig. 7B). While for ORFB, ORFC, and ORFD no function could be hypothesized, ORFA showed low conservation to known RdRp palm domains from different viruses when

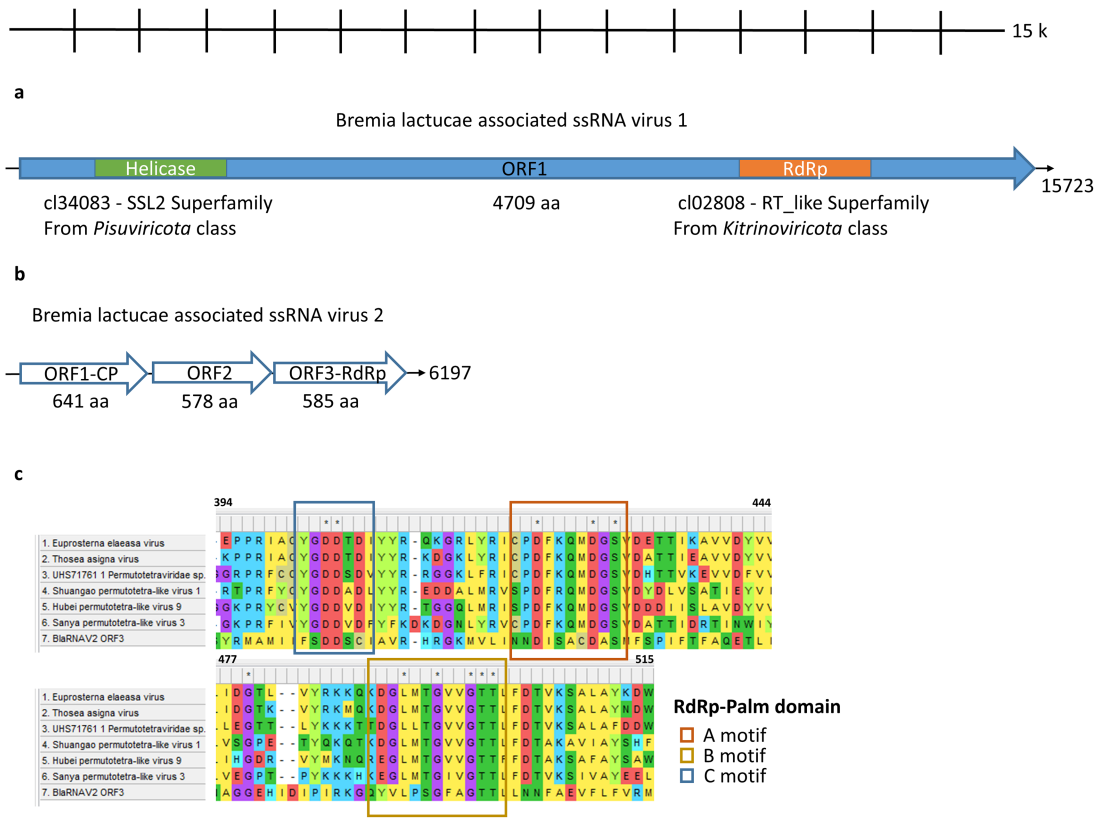


Figure 5. Schematic representation of the positive sense single-stranded RNA viruses and alignment between permuted RdRps. Panel (A) shows a schematic representation of Bremia lactucae-associated ssRNA virus 1 and panel (B) shows a representation of the genome of Bremia lactucae-associated ssRNA virus 2. Thin arrows represent the genomic segments; larger arrows represent the putative ORFs. Panel (C) shows a portion of the alignment between BlaRNAV2 RdRp and some permutotetraviruses. Numbers above the portions show the position on the total alignment. The portion shown contains the three motifs (highlighted with square boxes) typical of the permuted RdRps (C-A-B).

investigated through Swiss-model platform (but not through blast similarity searches). Indeed, based on the folding prediction for ORFA and for the RdRp selected as a model (with the QMEANDisCo value of 0.36 ± 0.05, namely the RdRp of the classical swine fever virus -CSFV- related to the family *Flaviviridae*), the alignment of the palm domain conserved motifs A, B, and C of the two sequences could be observed (Supplementary Fig. S9). Because of the novelty of this virus, we confirmed the *in silico* assembly through overlapping PCR and sanger sequencing of the fragments (Supplementary Fig. S10). Taken together, these results suggest that BlaORFanV1 represents a new taxa of RNA virus that is currently unrelated to known viruses, making it impossible to reliably associate it to known groups in the *Riboviria*. As a further proof of this taxonomic uncertainty, we blasted the RdRp of BlaORFanV1 against the Riboviria database (Neri et al. 2022) obtaining hits against RdRps classified in different phyla including *Pisuviricota*, *Kitrinoviricota*, *Duplornaviricota*, and RdRps for which the taxonomic classification was impossible. Interestingly, the closest hit was with an RdRp (ND_158467) with an identity percentage of 28.40 per cent on 40 per cent of the query cover (evalue 9e-67) that the authors could not classify reliably showing how distant these two proteins are and further demonstrating the difficulties in placing BlaORFanV1 putative RdRp in the context of the current viral RdRp phylogeny.

Discussion

The investigation of the viruses infecting *B. lactucae* allowed, for the first time, the association between RNA viruses and this important plant pathogen, also contributing to better understanding the viral diversity associated with oomycetes belonging to *Peronosporales*. Interestingly, only two viruses found in this study were distinct isolates of viruses already reported in literature (the partitivirus Bremia lactucae associated partitivirus 1 and the splipalmivirus Bremia lactucae associated splipalmivirus 1); in both cases, the viruses were found in oomycetes-related studies focused on *Plasmopara viticola*, supporting a specificity for oomycetes hosts for these two viruses.

Results collected in this study further confirm that eukaryotic microorganisms (including fungi and oomycetes) are important reservoirs of viral diversity from an evolutionary point of view. Indeed, we could not only describe new mycoviral lineages, thanks to homology-based searches and ORFan investigations, but we could also report evidence of structural or replicative reorganization of RNA viruses helpful to understand how they evolve. Both these phenomena could be reported for Bremia lactucae associated ditolivirus 1; in this case, the bi-segmentation of the viral genome seems to originate from a division of a monopartite ancestor virus rather than from the acquisition of a second genomic fragment from different origin. Furthermore, in PvlaTlV5, a virus from the same clade (possibly a new family in the Duplornaviricota) but detected in the oomycetes *P. viticola*, we show evidence of a very common evolutionary strategy: extant repurposing of a gene function after duplication, which, together with exaptation from hosts, is one of the main motors of virus evolution (Koonin, Dolja, and Krupovic 2022).

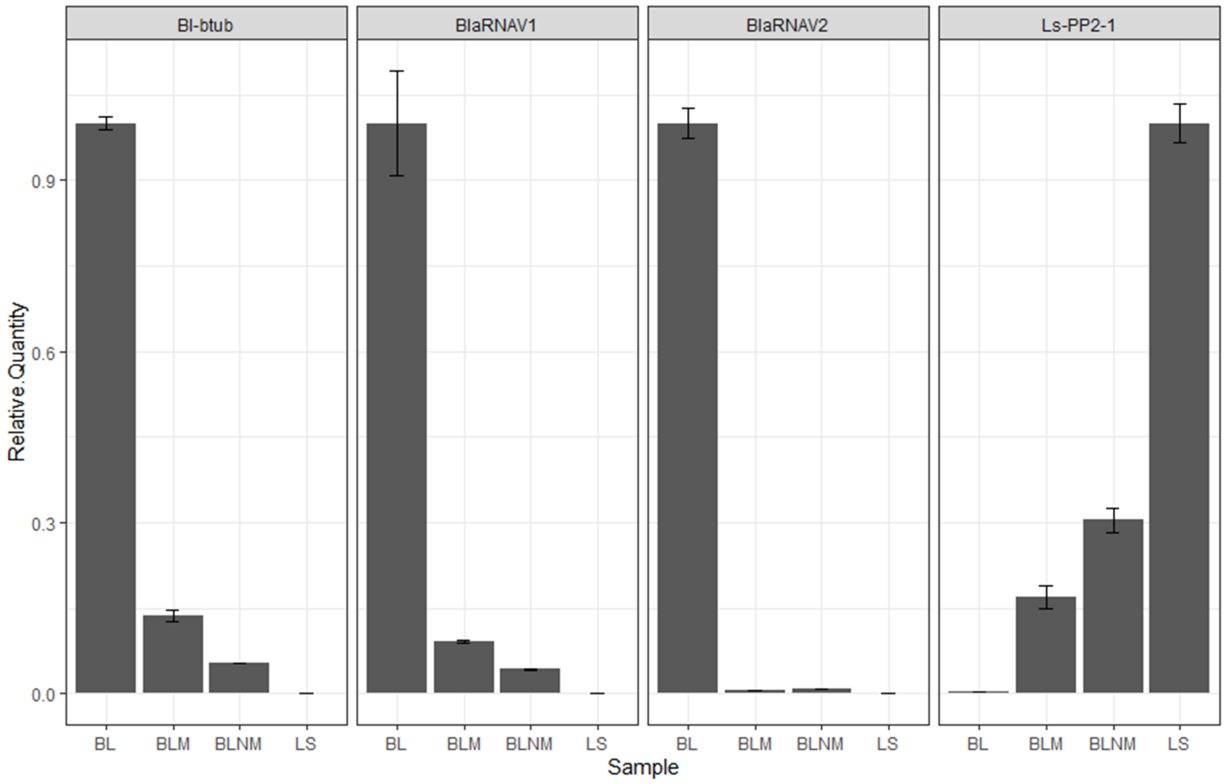


Figure 6. Relative abundance distribution of BlaRNAV1 and 2 in fractions of lettuce infected with *Bremia lactucae* isolate BL20EU. The picture shows the relative abundance of BlaRNAV1, BlaRNAV2, the beta-tubulin gene from *Bremia lactucae* and the PP2-1 gene from *Lactuca sativa* obtained through qRT-PCR. Each target was tested on different samples: BL is the RNA from an enrichment of *Bremia lactucae* zoosporangia isolated from infected lettuce seedlings. BLM is the RNA extracted from *B. lactucae* infected seedlings with *B. lactucae* zoosporangia clearly visible on the seedling surface. BLNM is the RNA extracted from *B. lactucae* infected lettuce in leaf areas without zoosporangia emerging from the seedling surface. LS is the RNA extracted from a lettuce seedling not inoculated with *B. lactucae*. The error bars show the standard error calculated on three replicates. The experiment was conducted with the *B. lactucae* isolate BL20EU, infected with both BlaRNAV1 and 2. Details on primers used for this experiment and purification protocol for the zoosporangia are available in the method section.

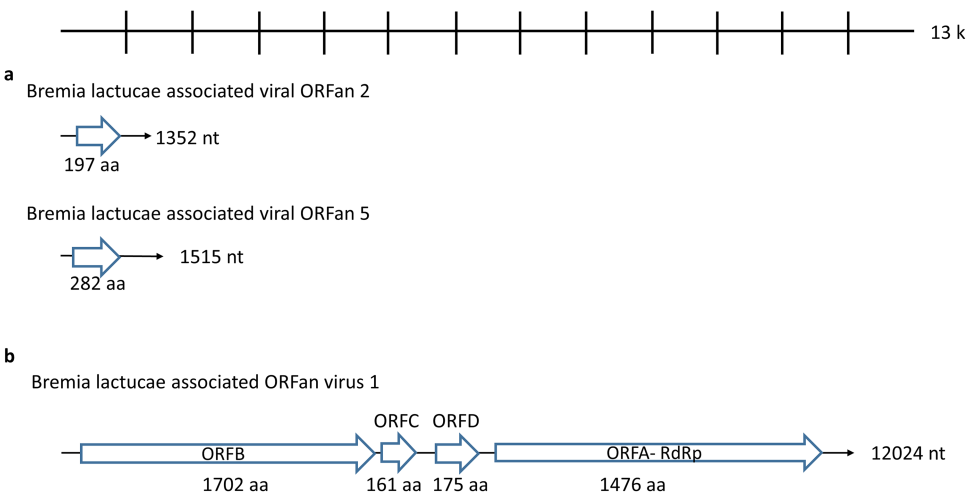


Figure 7. Schematic representation of the viral ORFan contigs. Thin arrows represent the genomic segments; larger arrows represent the putative encoded proteins.

The emergence of multipartitism is the focus of many studies that try to explain the condition required for the occurrence of this event; in particular, some models explain the multipartitization of a monopartite virus with the improved ability of the fragmented virus to modulate the production of the different viral proteins carried on multiple segments in response to variations of the environmental conditions (Zwart and Elena 2020). Other studies link multipartitism to a fitness benefit for the fragmented virus, leading to smaller and more stable viral particles hosting only one genome segment compared to the larger virions hosting the monopartite counterpart (Ojosnegros et al. 2011). Finally, a recent study sets up a model to explain multipartitism stating

that fragmented version of the viral genome (carrying fundamental genes for the virus) could emerge without any selective pressure from the environment, simply from the production of what the authors call 'cheaters' (like defective interfering RNAs). Such elements can complement each other and, in the presence of multiple fragments carrying the necessary genes, can complete the replication cycle (Leeks et al. 2023). It is interesting that, in this model, authors suggest that the possibility to have a multipartitization event is correlated with the production of cheats (really frequent for RNA viruses), and that the phenomenon is more frequent in chronic infection as observed in plant; these are exactly the condition observed for RNA mycoviruses, where the virus is spread vertically through spores and horizontally through vegetative compatibility between mycelia.

The stranded mapping data also show, for BlaDTV1 and other candidates for the proposed group ditolivirus, the unusual accumulation of more negative strands of the genome. This was already observed in other dsRNA viruses, and could be linked to an adaptation to the host antiviral system (Dave et al. 2019; Chiapello et al. 2020); moreover, the accumulation of more negative sense strand was also reported in other groups of putatively naked RNA mycoviruses (as the splipalmivirus Oidiodendron maius splipalmivirus 1 and a specific clade of ORFan viruses called gammaormycoviruses) (Sutela et al. 2020; Forgia et al. 2022), making fungi (and other mycelia-like growing microorganisms like oomycetes) a suitable source for the emergence of a negative sense-oriented viral lifestyle. Indeed, current theories for the emergence of negative sense RNA viruses propose their origin from dsRNA viruses related to the Phylum *Duplornaviricota* (Koonin et al. 2020), but in a cellular context where viruses are easily losing their capsid proteins and replicate as naked RNA with multiple coinfection which may elicit phenomena of acquisition of nucleocapsids from different sources (either from host or from other viral taxa) (Koonin, Dolja, and Krupovic et al. 2022). Our work also supports a newer topology for the phylogenetic tree of Riboviria, where the negative strand RNA viruses (Phylum Negarnaviricota) are more basal and are not derived from dsRNA (Neri et al. 2022); the authors themselves seem to suggest a possible artifact, since the authors associated minus strand RNA viruses mostly to insects, plants, and animals, which would be inconsistent with an early rise of this replication strategy; in recent years, a great diversity of negative strand viruses in both fungi and oomycetes (Liu et al. 2014; Botella et al. 2020, 2022; Pagnoni et al. 2023) (including the new Yue-like viruses we describe here) was uncovered making this new tree topology also plausible.

Differently from BlaDTV1, the origin of Bremia lactucae-associated ssRNA virus 1 and 2 seems related to a horizontal gene transfer (HGT) event between distantly related viral genomes. BlaRNAV1 shows conserved helicase and RdRp domains with distant origin (the RdRp domain is related to the family *Flaviviridae* while the helicase domain is related to the family *Potyviridae*); this kind of event was already observed in other mycoviruses and reported in literature, as the case of the large hypovirus detected in *Rhizoctonia solani* and called Rhizoctonia solani hypovirus 2, where a helicase domain related to the order *Nidovirales* could be found in the polyprotein (Abdoulaye et al. 2021). In the case of Rhizoctonia solani hypovirus 2, the acquisition of the distantly related helicase from a HGT event was explained with the need of an adapted helicase to manage the large size of the hypovirus genome. Since BlaRNAV1 also displays a large genome of over 15 kb, it is possible that the potyvirid helicase could have been acquired to manage the same function.

An HGT event could also be supposed for BlaRNAV2, where a tombusvirus-related coat protein is associated to a permuted RdRp that could not be linked reliably to other viral taxonomic groups. Permuted RdRps could be found in phylogenetically unrelated RNA plant viruses in the order *Tymovirales* (Sabanadzovic, Ghanem-Sabanadzovic, and Gorbalenya 2009) and in the family *Permutotetraviridae*, and a larger group of arthropod associated viruses whose numbers are increasing thanks to the latest metatranscriptomic works finding new elements (Wolf et al. 2020; Chen et al. 2022). Even if BlaRNAV2 RdRp was detected through comparison with the folded structure of a permutotetravirus, no conservation could be observed at the sequence level between BlaRNAV2- encoded RdRp and that of known groups of permuted viruses. Thus, BlaRNAV2 seems to represent a completely new group supporting the evidence of multiple independent origins of the permuted RdRp as described previously (Wolf et al. 2020).

The analysis of the ORFan sequences resulting from the metatranscriptome was crucial to complete the characterization of the new group of tripartite yue-like viruses associated with oomycetes, and to detect a new viral ORFan encoding for a low conserved RdRp. Indeed, in the case of the yue-like viruses related to Bremia lactucae associated yue-like virus 2, RdRp encoding segments could be detected through homology approach in recent works (Chiapello et al. 2020), but the putative complete genome could be detected only through investigating the ORFan sequences, adding further segments that encode proteins with undetectable similarities to those existing in databases. The putative proteins encoded by BlaYlV2 RNA2 and RNA3 were sufficiently conserved to detect other genomic fragments associated to the yue-like virus related to BlaYlV2, contributing to a better comprehension of the genome organization of at least one of the group of viruses related to yueviruses that are still poorly characterized. To our knowledge, this is currently the first available biological sample that can be used to further our knowledge on this recently detected group of negative sense RNA viruses.

Finally, the detection of Bremia lactucae-associated ORFan virus 1 will contribute to expansion of the detection of new RNA viruses, thanks to the availability in public databases of newly detected and annotated RdRp sequences with low conservation to viruses already reported. This work also reports two viral fragments (BlaVORFan2 and BlaVORFan5) for which an association with an RdRp is still unknown, leaving an open question about their nature as putative viral fragment: it will be interesting to further investigate the conserved structural domain shared by this RNA-only element and two ORFan proteins encoded by the genomes of *B. lactucae* and *P. halstedii*; this raises also questions about possible horizontal transfer events between the virus and its host, involving protein domains respectively encoded by putative RNA-only elements and genomic loci.

In conclusion, the first identification of viruses like BlaORFanV1, BlaRNAV1, BlaRNAV2, and BlaYlV2 supports the hypothesis that some viral taxa are often enriched or specific in oomycetes as was reported for the bunyaviruses detected in *Phytophthora* and *Halophytophthora* works (Botella et al. 2020; Poimala et al. 2021), showing how our knowledge on oomycetes-associated viruses is still poor. Moreover, this study also proposes for the first time a group of dsRNA viruses, likely specific for oomycetes, that was tentatively named ditolivirus.

So far, a single word has been used to define viruses that infect both fungi and oomycetes: this has historical reasons, since the first viruses identified in oomycetes were not specific for this clade but were similar to those infecting fungi. Now, the overall

diversity of viruses infecting oomycetes and fungi shows more specificity and we suggest that a new word should be used to describe viruses that infect oomycetes (e.g. ooviruses or oomycovirus); in fact, oomycetes share with fungi ecological niches, filamentous growth, and reproduction through spore formation, but their diploid vegetative ifae, lack of septa, their partial sterol autotrophy, and absence (or very little presence) of chitin in their cell wall sets them apart from true fungi; in this respect, even the most recent phylogenomic classifications of eukaryotes confirm the groundbreaking findings of the end of the last century (Baldauf et al. 2000) that showed that oomycetes are close relatives of heterokont golden-brown algae placing them in the Telonemia Stramenopiles Alveolata Rhizaria supergroup (that includes stramenopiles) while fungi (and animals) are in the Amorphea supergroup that includes opisthokonta (Burki et al. 2020) (therefore placing fungi and oomycetes in distinct kingdoms). The distinct evolutionary history of their relationship with viruses is therefore not surprising, with some occasional evidence of horizontal transfers due to their similar lifestyles and shared spatial and ecological niches. The role of fungi as source of viruses infecting plants is supported by many lines of evidence (Dolja, Krupovic, and Koonin 2020); a comparative in-depth analysis establishing the role of viruses infecting oomycetes remains to be carried out, but our finding of a partitivirus and of an endornavirus in *B. lactucae* suggest cross-kingdom horizontal transfer occurring also between oomycetes and plants.

Materials and methods

B. lactucae maintenance

Eleven isolates of *Bremia lactucae* were retrieved from the Naktuinbouw collection (www.naktuinbouw.com) that are characterized for their capacity to overcome different resistance genes in lettuce and maintained on lettuce seedlings. Accessions used for this work are listed in Table 2 with the official code names from the original collection. *Lactuca sativa* seedlings (cultivar green tower) were put on a layer of cotton inside a closed transparent plastic box each Tuesday of every week and hydrated with water; lettuce boxes were maintained in a growth chamber at 13°C with a light/dark cycle of 12 hours. After nine days, germinated seedlings were treated with a water suspension containing *B. lactucae* zoosporangia collected from previously infected seedlings. The zoosporangia collection was made by placing infected lettuce seedlings (with clearly visible *B. lactucae* sporulation on the surface) in a 50 mL falcon tube and adding deionized water. The tube was shaken for a few seconds and filtered through a clean gauze to remove the plant tissue. Suspension was then sprayed on fresh seedlings or used for RNA extraction. Zoosporangia suspensions could be stored at –20°C for a few weeks as a backup inoculum to restart the cultures if necessary.

RNA extraction and HTS sequencing

Total RNA was extracted from zoosporangia suspension collected as explained above. Suspensions were put in a 2-mL O-ring tube and zoosporangia were precipitated using a centrifuge. Most of the liquid was removed by pipetting without touching the zoosporangia pellet at the bottom of the tube. Glass beads (0.5 mm diameter) were then added to the tube to break the cellular tissue through a MP Biomedicals™ FastPrep-24™ (Thermo Fisher Scientific, Waltham, MA, USA). Zoosporangia lysate was then extracted using Spectrum™ Plant Total RNA Kit (Merck, Darmstadt, Germany) following the suggested protocol. RNAs were quantified with a NanoDrop 2000 apparatus (Thermo Fisher Scientific, Waltham, MA, USA) and each of them were pooled in one HTS sample containing the same amount of RNA from each of the single isolates. Library preparation and sequencing through Illumina platform were performed by Macrogen Europe (Amsterdam, The Netherlands) performing rRNA depletion through TruSeq Stranded Total RNA Ribo Zero Human/Mouse Gold kit and a NovaSeq 6000 platform. Reads obtained from this work are deposited with the following bioproject ID: PRJNA891641.

Bioinformatic analysis

Resulting reads were cleaned from low-quality scores, adapters, and artifacts using bbmap and assembled using Trinity (Haas et al. 2013) software version 2.9.1. Virome analysis was performed through homology approach and ORFan investigation. Homology search was performed using Diamond software (Buchfink, Xie, and Huson 2015) and blasting the Trinity assembly against the NCBI nr database through tblastn function. Viral contigs were selected from the blast output filtering the contigs blasting against the viral accession and checking them manually before the detection through RT-PCR in the single oomycetes isolates. ORFan investigation was performed as detailed previously (Forgia et al. 2022) through selection of the contigs longer than 1,000 bp and encoding for proteins of at least 150 amino acids that don't show hits against the NCBI nr database. All the ORFan contigs selected were mapped to retrieve the number of reads mapping on each sense of the sequences and only the one showing reads on both senses (indication of possible dsRNA accumulation, typical of viral replication) were further characterized through RT-PCR to confirm their presence only as RNA sequences. Read mappings were performed with bowtie2 (Langmead and Salzberg 2012) software and read strand information was extracted from the bowtie2 output through Samtools (Li et al. 2009). ORF prediction was performed using ORFinder (Rombel et al. 2002) and conserved protein domains were searched through a conserved domain search tool from NCBI. Structural analysis performed for BlaRNAV2 was made through the Swiss-model browser platform (Waterhouse et al. 2018). Nucleotide alignments were performed to display end conservation, and pairwise comparisons were made through the 'global align' option implemented in the blast interface. When comparing the ends from more than two viral fragments, the clustal omega platform (Sievers et al. 2011) was used from the browser. BlaVORFan2 and BlaVORFan5 3D protein structure was predicted in silico with Alphafold2 (Jumper et al. 2021) implemented using ColabFold v1.5.2 (Mirdita et al. 2022). Obtained models were used for database searches of 3D structural models using DALI server (Holm 2020) and Foldseek search server (van Kempen et al. 2023). Obtained models were then employed for structural comparison and Root Mean Square Deviation (RMSD) estimation using UCSF ChimeraX 'matchmaking' function, with default parameters (Pettersen et al. 2021).

RT-PCR analysis and viral distribution

Complementary DNAs were synthetized from total RNA samples using High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, USA) for real-time applications, and RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA) for overlapping PCR applications; all the cDNAs were diluted 1:5 prior use. A list of primers used for the work is available in Supplementary Table S5; all the primers used were designed using Primer3 (Untergasser et al. 2012). Overlapping PCR was performed using OneTaq DNA polymerases (New England Biolabs, Ipswich, MA, USA) while qRT-PCR was performed using a

Bio-Rad CFX apparatus and 2X Syber green mix in 10 μl of volume per sample (Bio-Rad, Hercules, CA, USA). To demonstrate the association between *B. lactucae* and BlaRNAV1, 2, and BlaVORFan2, total RNA was extracted from uninfected lettuce seedlings, infected lettuce seedling from areas without clear sporulation on the surface, and infected lettuce seedling from areas with visible zoosporangia on the surface, and finally from zoosporangia isolated as described before. The same amount of RNA (around 1 μg) was DNAsed using DNAse 1 (Thermo Fisher Scientific, Waltham, MA, USA) and retrotranscribed for real-time qRT-PCR application. Each sample was tested in triplicate with primers detecting a reference gene for lettuce and for *B. lactucae* as a measure of the proportion between the two organisms in the sample (*PP2-1* for *Lactuca sativa* and *β-tubulin* gene for *B. lactucae*); at the same time, BlaRNAV1 and 2 were tested in triplicate in all the samples. The mean and standard error of the normalized ΔCt values were plotted to show the viral distribution across the samples using the ggplot2 library on R.

RACE analysis

RACE analysis was performed on Bremia lactucae associated viral ORFan 2 to obtain the complete sequence. The protocol used (Rastgou et al. 2009) consisted in the production of cDNA using specific primers on the 5′ and 3′ (Supplementary Table S5) of the contig using Superscript IV Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA). Obtained cDNAs were tagged with a polyA or polyG using Terminal Deoxynucleotidyl Transferase (Promega, Madison, WI, USA), and polyT or polyC primers were used to amplify PCR fragments coupling them with specific primers (Supplementary Table S5). Obtained PCR bands were cloned in pGEMT-easy vector (Promega, Madison, WI, USA) and sequenced to obtain at least three clones confirming the end sequence.

Phylogenetic analysis and identity matrix

Phylogenetic analyses were performed through alignment of RdRp sequences between different viral groups that included the closest hits by blast and some representative of close taxonomically characterized clades. RdRp sequences were aligned through Multiple Alignment using Fast Fourier Transform (MAFFT) (Katoh, Rozewicki, and Yamada 2019) and obtained multiple alignments were submitted to IqTree webserver (Trifinopoulos et al. 2016) for phylogenetic tree construction through maximum likelihood method. The model selection was performed by ModelFinder implemented in IqTree (Kalyaanamoorthy et al. 2017), and ultrafast bootstrap analysis (Hoang et al. 2018) was performed with 1,000 replicates. The lists of RdRp included in each phylogenetic tree, together with the NCBI accession numbers, are found in Supplementary Table S6. Identity matrices were produced starting from a MAFFT alignment of the RdRp encoded by the fusagraviruses. Resulting alignment was submitted to Discovery studio to produce the identity matrix. The obtained identity matrix was used to generate the picture using the libraries reshape2, tidyverse, and ggplot2 on R.

Data availability

Raw reads from the illumina sequencing of the *B. lactucae* pooled isolate are available on the NCBI database with the bioproject ID: PRJNA891641. The viral genomes characterized in the work are available in the NCBI with the accession numbers listed in Table 1 for each viral genome segment or viral ORFan.

Supplementary data

Supplementary data is available at *Virus Evolution* Journal online.

Acknowledgements

European Union's Horizon 2020 research and innovation program (773567). The authors would like to thank Elena Zocca for providing technical assistance with lettuce plant growth.

Conflict of interest: Authors declare no conflict of interest.

References

Abdoulaye, A. H. et al. (2021) 'Two Distant Helicases in One Mycovirus: Evidence of Horizontal Gene Transfer between Mycoviruses, Coronaviruses and Other Nidoviruses', *Virus Evolution*, 7: veab043.

Aihara, M. et al. (2018) 'Infection by Magnaporthe Oryzae Chrysovirus 1 Strain A Triggers Reduced Virulence and Pathogenic Race Conversion of Its Host Fungus, *Magnaporthe Oryzae*', *Journal of General Plant Pathology*, 84: 92–103.

Baldauf, S. L. et al. (2000) 'A Kingdom-level Phylogeny of Eukaryotes Based on Combined Protein Data', *Science*, 290: 972–7.

Botella, L. et al. (2020) 'Marine Oomycetes of the Genus *Halophytophthora* Harbor Viruses Related to Bunyaviruses', *Frontiers in Microbiology*, 11: 1467.

Botella, L. et al. (2022) 'Natural Populations from the *Phytophthora Palustris* Complex Show a High Diversity and Abundance of ssRNA and dsRNA Viruses', *Journal of Fungi*, 8: 1118.

Botella, L., and Jung, T. (2021) 'Multiple Viral Infections Detected in *Phytophthora Condilina* by Total and Small RNA Sequencing', *Viruses*, 13: 620.

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

Burki, F. et al. (2020) 'The New Tree of Eukaryotes', *Trends in Ecology and Evolution*, 35: 43–55.

Cai, G. H., and Hillman, B. I. (2013) 'Phytophthora Viruses', in S. A. Ghabrial (ed.) *Advances in Virus Research, Vol 86: Mycoviruses (Advances in Virus Research, 86)*, pp. 327–50. Cambridge, MA, USA: Academic Press.

Chen, Y.-M. et al. (2022) 'RNA Viromes from Terrestrial Sites across China Expand Environmental Viral Diversity', *Nature Microbiology*, 7: 1312–23.

Chiapello, M. et al. (2020) 'Analysis of the Virome Associated to Grapevine Downy Mildew Lesions Reveals New Mycovirus Lineages', *Virus Evolution*, 6: veaa058.

Chiba, Y. et al. (2021) 'Discovery of Divided RdRp Sequences and a Hitherto Unknown Genomic Complexity in Fungal Viruses', *Virus Evolution*, 7: veaa101.

Das, S. et al. (2021) 'Proof of Concept of the Yadokari Nature: A Capsidless Replicase-Encoding but Replication-Dependent Positive-Sense Single-Stranded RNA Virus Hosted by an Unrelated Double-Stranded RNA Virus', *Journal of Virology*, 95: 10–1128.

Dave, P. et al. (2019) 'Strand-specific Affinity of Host Factor hnRNP C1/C2 Guides Positive to Negative-strand Ratio in Coxsackievirus B3 Infection', *RNA Biology*, 16: 1286–99.

Dolja, V. V., Krupovic, M., and Koonin, E. V. (2020) 'Deep Roots and Splendid Boughs of the Global Plant Virome', *Annual Review of Phytopathology*, 58: 23–53.

Edgar, R. C. et al. (2022) 'Petabase-scale Sequence Alignment Catalyses Viral Discovery', *Nature*, 602: 142–7.

Espino-Vazquez, A. N. et al. (2020) 'Narnaviruses: Novel Players in Fungal-bacterial Symbioses', *The ISME Journal*, 14: 1743–54.

Forgia, M. et al. (2021) 'Virome Characterization of *Cryphonectria Parasitica* Isolates from Azerbaijan Unveiled a New Mymonavirus and a Putative New RNA Virus Unrelated to Described Viral Sequences', *Virology*, 553: 51–61.

Forgia, M. et al. (2022) 'Three New Clades of Putative Viral RNA-Dependent RNA Polymerases with Rare or Unique Catalytic Triads Discovered in Libraries of ORFans from Powdery Mildews and the Yeast of Oenological Interest *Starmerella Bacillaris*', *Virus Evolution*, 8: veac038.

Haas, B. J. et al. (2013) 'De Novo Transcript Sequence Reconstruction from RNA-seq Using the Trinity Platform for Reference Generation and Analysis', *Nature Protocols*, 8: 1494–512.

Heller-Dohmen, M. et al. (2011) 'The Nucleotide Sequence and Genome Organization of *Plasmopara Halstedii* Virus', *Virology Journal*, 8: 1–8.

Hoang, D. T. et al. (2018) 'UFBoot2: Improving the Ultrafast Bootstrap Approximation', *Molecular Biology and Evolution*, 35: 518–22.

Holm, L. (2020) 'Using Dali for Protein Structure Comparison', in Gaspari, Z. (ed.) *Structural Bioinformatics: Methods and Protocols (Methods in Molecular Biology, 2112)*, pp. 29–42. Humana Press: New York, NY, USA.

Jumper, J. et al. (2021) 'Highly Accurate Protein Structure Prediction with AlphaFold', *Nature*, 596: 583–9.

Kalyaanamoorthy, S. et al. (2017) 'ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates', *Nature Methods*, 14: 587–9.

Kanhayuwa, L. et al. (2015) 'A Novel Mycovirus from *Aspergillus Fumigatus* Contains Four Unique dsRNAs as Its Genome and Is Infectious as dsRNA', *Proceedings of the National Academy of Sciences of the United States of America*, 112: 9100–5.

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019) 'MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization', *Briefings in Bioinformatics*, 20: 1160–6.

Koonin, E. V. et al. (2020) 'Global Organization and Proposed Megataxonomy of the Virus World', *Microbiology and Molecular Biology Reviews*, 84: e00061–19.

Koonin, E. V., Dolja, V.V., and Krupovic, M. (2022) 'The Logic of Virus Evolution', *Cell Host & Microbe*, 30: 917–29.

Langmead, B., and Salzberg, S. L. (2012) 'Fast Gapped-read Alignment with Bowtie 2', *Nature Methods*, 9: 357–9.

Leeks, A. et al. (2023) 'Cheating Leads to the Evolution of Multipartite Viruses', *PLoS Biology*, 21: e3002092.

Li, H. et al. (2009) 'The Sequence Alignment/map Format and SAMtools', *Bioinformatics*, 25: 2078–9.

Liu, L. et al. (2014) 'Fungal Negative-stranded RNA Virus that Is Related Tobornaviruses and Nyaviruses', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 12205–10.

Marzano, S. L. et al. (2016) 'Identification of Diverse Mycoviruses through Metatranscriptomics Characterization of the Viromes of Five Major Fungal Plant Pathogens', *Journal of Virology*, 90: 6846–63.

Michelmore, R., and Wong, J. (2008) 'Classical and Molecular Genetics of *Bremia Lactucae*, Cause of Lettuce Downy Mildew', *European Journal of Plant Pathology*, 122: 19–30.

Mirdita, M. et al. (2022) 'ColabFold: making protein folding accessible to all', *Nat Methods*, 19: 679–82.

Neri, U. et al. (2022) 'Expansion of the Global RNA Virome Reveals Diverse Clades of Bacteriophages', *Cell*, 185: 4023–37.

Nerva, L. et al. (2017) 'Transmission of Penicillium Aurantiogriseum Partiti-like Virus 1 to a New Fungal Host (*Cryphonectria Parasitica*) Confers Higher Resistance to Salinity and Reveals Adaptive Genomic Changes', *Environmental Microbiology*, 19: 4480–92.

Nerva, L. et al. (2018) 'Mycoviruses Mediate Mycotoxin Regulation in *Aspergillus Ochraceus*', *Environmental Microbiology*, 21: 1957–68.

Ojosnegros, S. et al. (2011) 'Viral Genome Segmentation Can Result from a Trade-Off between Genetic Content and Particle Stability', *PLoS Genetics*, 7: e1001344.

Pagnoni, S. et al. (2023) 'A Collection of Trichoderma Isolates from Natural Environments in Sardinia Reveals A Complex Virome that Includes Negative-sense Fungal Viruses with Unprecedented Genome Organizations', *Virus Evolution*, 9: vead042.

Parra, L. et al. (2016) 'Rationalization of Genes for Resistance to *Bremia Lactucae* in Lettuce', *Euphytica*, 210: 309–26.

Pettersen, E. F. et al. (2021) 'UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers', *Protein Science*, 30: 70–82.

Picarelli, M. A. S. C. et al. (2019) 'Extreme Diversity of Mycoviruses Present in Isolates of Rhizoctonia Solani AG2-2 LP from *Zoysia Japonica* from Brazil', *Frontiers in Cellular & Infection Microbiology*, 9: 244.

Poimala, A. et al. (2021) 'Viral Diversity in *Phytophthora Cactorum* Population Infecting Strawberry', *Environmental Microbiology*, 23: 5200–21.

Poimala, A. et al. (2022) 'Bunyaviruses Affect Growth, Sporulation, and Elicitin Production in *Phytophthora Cactorum*', *Viruses*, 14: 2596.

Raco, M. et al. (2022) 'High Diversity of Novel Viruses in the Tree Pathogen *Phytophthora Castaneae* Revealed by High-Throughput Sequencing of Total and Small RNA', *Frontiers in Microbiology*, 13: 911474.

Rastgou, M. et al. (2009) 'Molecular Characterization of the Plant Virus Genus *Ourmiavirus* and Evidence of Inter-kingdom Reassortment of Viral Genome Segments as Its Possible Route of Origin', *Journal of General Virology*, 90: 2525–35.

Rigling, D., and Prospero, S. (2017) '*Cryphonectria Parasitica*, the Causal Agent of Chestnut Blight: Invasion History, Population Biology and Disease Control', *Molecular Plant Pathology*, 19: 7–20.

Rombel, I. T. et al. (2002) 'ORF-FINDER: A Vector for High-throughput Gene Identification', *Gene*, 282: 33–41.

Sabanadzovic, S., Abou Ghanem-Sabanadzovic, N., and Gorbalenya, A. E. (2009) 'Permutation of the Active Site of Putative RNA-dependent RNA Polymerase in a Newly Identified Species of Plant Alpha-like Virus', *Virology*, 394: 1–7.

Shi, M. et al. (2016) 'Redefining the Invertebrate RNA Virosphere', *Nature*, 540: 539–43.

Sievers, F. et al. (2011) 'Fast, Scalable Generation of High-quality Protein Multiple Sequence Alignments Using Clustal Omega', *Molecular Systems Biology*, 7: 539.

Sutela, S. et al. (2020) 'The Virome from a Collection of Endomycorrhizal Fungi Reveals New Viral Taxa with Unprecedented Genome Organization', *Virus Evolution*, 6: veaa076.

Trifinopoulos, J. et al. (2016) 'W-IQ-TREE: A Fast Online Phylogenetic Tool for Maximum Likelihood Analysis', *Nucleic Acids Research*, 44: W232–W35.

Turina, M. et al. (2018) 'The Virome of the Arbuscular Mycorrhizal Fungus *Gigaspora Margarita* Reveals the First Report of DNA Fragments Corresponding to Replicating Non-retroviral RNA Viruses in Fungi', *Environmental Microbiology*, 20: 2012–25.

Turina, M., and Rostagno, L. (2007) 'Virus-induced Hypovirulence in *Cryphonectria Parasitica*: Still an Unresolved Conundrum', *Journal of Plant Pathology*, 89: 165–78.

Untergasser, A. et al. (2012) 'Primer3-new Capabilities and Interfaces', *Nucleic Acids Research*, 40: e115.

Urayama, S. I., Takaki, Y., and Nunoura, T. (2016) 'FLDS: A Comprehensive dsRNA Sequencing Method for Intracellular RNA Virus Surveillance', *Microbes & Environments*, 31: 33–40.

van Kempen, M. et al. (2023) 'Fast and Accurate Protein Structure Search with Foldseek', *Nature Biotechnology*, 1–4.

Wang, L. et al. (2016) 'Two Novel Relative Double-Stranded RNA Mycoviruses Infecting *Fusarium Poae* Strain SX63', *International Journal of Molecular Sciences*, 17: 641.

Waterhouse, A. et al. (2018) 'SWISS-MODEL: Homology Modelling of Protein Structures and Complexes', *Nucleic Acids Research*, 46: W296–303.

Wolf, Y. I. et al. (2020) 'Doubling of the Known Set of RNA Viruses by Metagenomic Analysis of an Aquatic Virome', *Nature Microbiology*, 5: 1262–70.

Yokoi, T. et al. (1999) 'The Nucleotide Sequence and Genome Organization of *Sclerophthora Macrospora* Virus B', *Virology*, 264: 344–9.

Yokoi, T., Yamashita, S., and Hibi, T. (2003) 'The Nucleotide Sequence and Genome Organization of *Sclerophthora Macrospora* Virus A', *Virology*, 311: 394–9.

Zhang, H. et al. (2020) 'A 2-kb Mycovirus Converts A Pathogenic Fungus into A Beneficial Endophyte for Brassica Protection and Yield Enhancement', *Molecular Plant*, 13: 1420–33.

Zhou, L. et al. (2021) 'A Mycovirus Modulates the Endophytic and Pathogenic Traits of A Plant Associated Fungus', *The ISME Journal*, 15: 1893–906.

Zwart, M. P., and Elena, S. F. (2020) 'Modeling Multipartite Virus Evolution: The Genome Formula Facilitates Rapid Adaptation to Heterogeneous Environments†', *Virus Evolution*, 6: veaa022.

Virus Evolution, 2024, **10(1)**, 1–16
DOI: https://doi.org/10.1093/ve/veae003
Advance Access Publication 5 January 2024
Research Article