

Characterization of virome of *Plasmodiophora brassicae* in southwest China

yue deng

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

yunhui he

Fuyang Normal University

xiuying qian

Fuyang Normal University

yaoying yu

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

xiaoqin huang

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

xiaoxiang yang

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

zhongmei zhang

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

pei song

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

yunxi xian

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

yong liu

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

kang zhou

Fuyang Normal University

lei zhang

zhanglei19296@163.com

Institute of Plant Protection, Sichuan Academy of Agricultural Sciences

Research Article

Keywords: clubroot, *Plasmodiophora brassicae*, high-throughput sequencings, RNA virus, virus diversity, virus-host interaction

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10899099/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Plasmodiophora brassicae is an obligate biotrophic protist that causes clubroot disease in cruciferous crops, leading to severe yield losses worldwide. Despite its economic importance, the viral community associated with *P. brassicae* remains entirely unexplored. In this study, we collected clubroot samples from 16 different regions in Sichuan Province and Guizhou Province, China, and employed high-throughput sequencing to investigate the viral diversity in *P. brassicae* resting spores. A total of 29 viruses were identified, belonging to 13 viral families, including 15 positive-sense single-stranded RNA (+ ssRNA) viruses, eight negative-sense single-stranded RNA (-ssRNA) viruses, five double-stranded RNA (dsRNA) viruses, and one DNA virus. Among these, 14 viruses were identified as putative novel viruses showing less than 90% sequence identity with known viruses. The most widely distributed viruses were Cucumber mosaic virus (100%), Turnip yellows virus (91.7%), and Brassica yellows virus 1 (75%). Correlation analysis revealed that the relative abundance of *Plasmodiophora brassicae* partitivirus 1 (PbPV1) and Brassica virus 1_Nap was significantly positively correlated with clubroot disease incidence ($p < 0.05$). However, no significant correlation was observed between virus alpha diversity and environmental factors including disease incidence, longitude, latitude, and altitude. These findings significantly expand our understanding of the viral diversity associated with *P. brassicae* and provide a foundation for future studies on the potential roles of these viruses in clubroot disease epidemiology.

1. Introduction

Plasmodiophora brassicae is an important obligate biotrophic parasite within the protistan kingdom that causes clubroot disease in cruciferous crops, primarily affecting oilseed rape, Chinese cabbage, and radish. The pathogen induces abnormal proliferation and gall formation in root tissues, resulting in stunted growth, wilting, and severe yield reductions or even total crop failure [1]. In China, clubroot disease has spread across all major cruciferous crop production areas, posing a serious threat to the sustainable production of cruciferous vegetables [2]. The entire life cycle of *P. brassicae* is completed underground, where it biotrophically induces root tissue hyperplasia to obtain energy resources for sexual reproduction. Resting spores of *P. brassicae* released from decomposed galls can survive in soil for more than 10 years and germinate to cause infection during subsequent cropping seasons [3]. Therefore, the germination of *P. brassicae* resting spores is critical for clubroot disease development.

In recent years, the role of biological factors in plant disease epidemics has received increasing attention. Microbial communities that coexist with plants and pathogens—including fungi, bacteria, protists, and viruses—can influence plant disease epidemics through complex interactions [4]. Viruses can play important roles in disease epidemics by affecting the biological functions of their pathogen hosts [5]. Currently, research on virus-host interactions has been widely conducted in various plant pathogens; however, no viruses have been reported in *P. brassicae* to date. Viruses are the most abundant biological entities on Earth, existing in almost all ecosystems and playing critical roles in regulating host population dynamics, driving microbial evolution, and maintaining the balance of ecosystem material circulation and energy flow [6]. With the continuous development of metagenomic sequencing technology and bioinformatics analysis methods, research on unknown viruses in various ecosystems has been greatly promoted, revealing that our current understanding of viral diversity is far from comprehensive [7, 8].

As a protist, *P. brassicae* persists in soil for extended periods. The inability to culture protists axenically has made the study of viruses in these organisms particularly challenging. In recent years, high-throughput sequencing technologies have revealed diverse viral groups within protist communities [9]. For example, marine phytoplankton can be infected by various double-stranded DNA viruses such as *Chloroviruses* [10], and amoebae harbor various giant viruses of the genus *Mimivirus* [11]. Although no viruses have been reported in *P. brassicae* to date, viruses have been described in its close relative *Polymyxa graminis*. *P. graminis* is a soil-borne obligate parasite of gramineous plants with a life cycle similar to that of *P. brassicae*, and both can persist in soil as resting spores for extended periods. *P. graminis* has been reported to serve as a vector for more than 14 plant viruses, including Chinese wheat mosaic virus (CWMV) [12].

In this study, we collected clubroot samples from 16 different regions in Sichuan province and Guizhou province to detect diverse viruses in *P. brassicae* through high-throughput sequencing. We also evaluated the differences and associations in viral diversity and relative abundance across different geographic locations and disease incidence levels. The results of this study contribute to revealing the diversity of viruses associated with *P. brassicae* and their potential relationships with clubroot disease development.

2. Materials and Methods

2.1. Sample collection, resting spore purification and pathogenicity analysis

Clubroot samples were collected from different cities and counties in Sichuan Province and Guizhou Province during November to February of 2023–2024 (Table 1). Geographic coordinates, host plant species, and disease incidence information were recorded for each sampling site. In 2023, clubroot samples were collected from Chengdu, Leshan, Nanchong, and Kangding city in Sichuan Province. In 2024, clubroot samples were collected from 10 cities in Sichuan Province and Guizhou Province during a similar timeframe. Disease incidence was assessed in five fields per location using a five-point random sampling method, with 100 plants surveyed per field.

Table 1
Information of clubroot disease survey and sample collection from 2023 to 2024.

Name	Bio Sample	Site	Year	Host plant	Latitude	Longitude	Altitude	Disease incidence
LS23	S1-2023	Leshan, Sichuan	2023	<i>Brassica napus</i>	29.07	103.99	403.8	-
NC23	S1-2023	Nanchong, Sichuan	2023	<i>Brassica napus</i>	31.07	106.62	327.31	-
CZ23	S1-2023	Chengdu, Sichuan	2023	<i>Brassica napus</i>	30.69	103.66	554.48	0.88 ± 0.07
KD23	S1-2023	Kangding, Sichuan	2023	<i>Brassica napus</i> and <i>Brassica rapa</i>	30.01	101.95	2766.87	0.80 ± 0.09
GD	S2-2024	Guiding, Guizhou	2024	<i>Brassica napus</i>	26.48	107.16	991.5	0.63 ± 0.11
KD	S2-2024	Kangding, Sichuan	2024	<i>Brassica napus</i>	30.01	101.95	2766.87	0.90 ± 0.06
LD	S2-2024	Luding, Sichuan	2024	<i>Brassica napus</i>	29.70	102.08	1943	0.63 ± 0.11
Aba	S2-2024	Lixian, Sichuan	2024	<i>Brassica rapa</i> subsp. <i>pekinensis</i>	31.64	103.18	1959.95	0.34 ± 0.09
XC	S2-2024	Xichang, Sichuan	2024	<i>Brassica rapa</i> subsp. <i>pekinensis</i>	27.89	102.26	1510	0.20 ± 0.09
DZ	S2-2024	Dazhou, Sichuan	2024	<i>Brassica napus</i>	31.63	107.73	487.5	0.38 ± 0.20
Ldi	S2-2024	Mianyang, Sichuan	2024	<i>Brassica napus</i>	31.29	104.75	452.77	0.06 ± 0.05
Hdi	S2-2024	Mianyang, Sichuan	2024	<i>Brassica napus</i>	31.29	104.75	452.77	0.77 ± 0.15
DY	S2-2024	Deyang, Sichuan	2024	<i>Brassica napus</i>	31.29	104.60	511	0.42 ± 0.18
LS	S2-2024	Leshan, Sichuan	2024	<i>Brassica oleracea</i> var. <i>capitata</i>	29.07	103.99	403.8	0.66 ± 0.15
NC	S2-2024	Nanchong, Sichuan	2024	<i>Brassica napus</i>	31.07	106.62	327.31	0.44 ± 0.15
YB	S2-2024	Yibin, Sichuan	2024	<i>Brassica napus</i>	28.85	104.62	391.44	0.71 ± 0.14

Collected clubroot samples were surface-cleaned and stored at -20 °C. For each sampling location, 300 g galls of clubroot were weighed for the extraction and purification of *P. brassicae* resting spores for subsequent total RNA extraction. Resting spores were purified as follows methods [13]: the outer layer of each gall was removed with a knife, and galls were surface-sterilized with 75% ethanol for 1 min, transferred to 1% sodium hypochlorite solution for 2 min, and rinsed three times with sterile water. Sterilized galls were cut and homogenized in 300 mL of sterile water, and the homogenate was filtered through four layers of cheesecloth. The filtrate was centrifuged at 3500 g for 10 min, and the pellet was washed three times with sterile water. The pellet was then resuspended in 10 mL of 50% sucrose solution, incubated for 10 min, and centrifuged at 4000 g for 10 min. The supernatant was transferred to a new 50 mL tube, and the pellet was washed three additional times with sterile water to obtain purified resting spores for total RNA extraction.

Pathogenicity analysis of clubroot-diseased samples collected from 13 regions was performed under greenhouse (23°C, and 16 h light: 8 h dark). Rapeseed seeds (Cultivar: Chuanyou 83, *Brassica napus*) were sown in 4 × 8-well plug trays filled with sterilized nutrient soil. Seven days after seed germination, plants were inoculated with a resting spore suspension of *Plasmodiophora brassicae*. Galls with comparable disease severity were selected to prepare resting spore suspensions, whose concentration was uniformly adjusted to 1×10^7 spores/mL, with 10 mL inoculum applied per plant. For each sample, a total of 30 plants were inoculated, arranged into three biological replicates with 10 plants per replicate, and each replicate was placed in a separate plug tray. Disease incidence and Disease severity index of clubroot were calculated after 40 days of inoculation [14].

2.2. Total RNA extraction and high-throughput sequencing

Total RNA of *P. brassicae* resting spores was extracted using the Quick-RNA Fungal/Bacterial Microprep Kit (Zymo Research, Cat#R2010) following the manufacturer's protocol with modifications. Briefly, purified resting spore pellets were transferred into BashingBead lysis tubes containing lysis buffer supplemented with β -mercaptoethanol. Cells were disrupted by bead beating at 5.0 m/s for 40 s, followed by 2 min of cooling on ice, and the bead-beating step was repeated once. After filtration, RNA was bound to the spin column, and on-column DNase I digestion was performed to eliminate genomic DNA contamination. The purified RNA was washed and eluted with RNase-free water. RNA integrity and purity were evaluated using agarose gel electrophoresis and spectrophotometry.

Total RNA samples from the same region (2024 collection) were pooled at equal molar quantities to generate 16 RNA samples for meta-transcriptome sequencing on an Illumina NovaSeq 6000 platform performed by Novogene Co., Ltd. (Beijing, China). Quality control was performed when obtaining the raw data using Trimmomatic (v0.41) [15]. Paired reads were discarded when more than 20% of bases in a single-end read had quality scores ($sQ \leq 20$). PCR amplification duplicates and polyX sequences were then removed to obtain clean reads. To eliminate the influence of host genes on subsequent analysis, clean reads were aligned to the host genome using bowtie2 to remove ribosomal and host sequences [16]. Clean reads were de novo assembled using MEGAHIT (v1.1.2; default parameters: --presets meta-large --min-contig-len 300) to obtain virus-related contigs [17]. BWA (v0.7.17) was used to align clean reads to the assembly results, and read utilization rates were calculated [18].

Virus-related contigs were aligned with the National Center for Biotechnology Information (NCBI) GenBank database through BLASTx (<https://www.ncbi.nlm.nih.gov/>) using DIAMOND (v0.8.22) software (<https://github.com/bbuchfink/diamond/>) [19]. Sequences with identity of 90% or higher were identified as known viruses annotated in the GenBank database. Sequences with identity below 90% or unaligned sequences were further aligned with the non-redundant protein (NR) database. For multi-segment RNA viruses, individual segments were aligned and annotated separately, and the information for each virus strain was summarized and compiled by segment.

2.3. Viral genome structure analysis and phylogenetic analysis

The assembly sequences of putative viruses were analyzed using the NCBI Conserved Domain Search Service (CD Search) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and Open Reading Frame Finder (ORF Finder) (<https://www.ncbi.nlm.nih.gov/orffinder/>) to identify viral ORFs and conserved domain. Based on the assembled viral sequence lengths, predicted ORFs, and conserved domain sequence information, genome structure maps of putative viruses were drawn using a uniform scale.

Sequence alignment was performed using BLASTn to retrieve complete or partial amino acid sequences (RNA dependent RNA polymerase or ORF) with high similarity to target RNA viruses from the NCBI public database (NCBI Virus) for phylogenetic analysis. Phylogenetic trees were constructed using MEGA 12 (v11.0) with the neighbor-joining algorithm, and branch support was evaluated through 1000 bootstrap replicates.

2.4. Virus abundance calculation and correlation analysis with environmental factors

Salmon (v1.9.0) software was used to map sequence reads to viral contigs [20]. Based on the results of virus sequence identification and taxonomic annotation, the abundance of putative RNA virus sequences was calculated. Alignment results with alignment lengths below 80% of the total read length were filtered out, and the proportion of viral reads was calculated. Transcripts per million (TPM) values were calculated for each viral contig. Based on the annotation results of viral contigs, the distribution of viral reads was summarized, and viruses with high abundance (at the species level) were identified.

Viral relative abundance was then calculated as the proportion of viral reads divided by the total reads in the library [21, 22]. The richness index was calculated using the *vegan* (v2.6-2) package in R [23]. The t-test was performed using the *stats* (v4.3.1) package in R (v4.4.3), and $p < 0.05$ was considered to indicate statistical significance. All data visualization was performed using *ggplot2* package in the R. Sankey diagrams illustrating virus species and taxonomic status were generated using the *networkD3* package. Heatmaps of relative abundance for selected viruses across samples were generated using the *heatmap* package. Correlation analysis between viral relative abundance and environmental factors (disease incidence, longitude, latitude, and altitude) was visualized using bubble plots generated with *ggplot2* package. Correlation analysis between virus alpha diversity indices and environmental factors was also visualized using *ggplot2* package.

3. Results

3.1. Virus diversity within resting spores of *P. brassicae*

A total of 16 clubroot-diseased samples were collected from 13 distinct locations in Sichuan and Guizhou Provinces during 2023–2024 (Fig. 1A and 1B). Results from pot assays of the same batch demonstrated that samples from all 13 locations retained normal infectivity (DI > 60%). Analysis of disease severity index revealed that the HDi sample from Mianyang and the XC sample from Xichang exhibited the strongest pathogenicity, whereas *P. brassicae* samples from ABa, DZ and GD displayed weaker pathogenicity compared with samples from other locations (Fig. 1C).

A total of 29 putative viruses were identified in all *P. brassicae* samples through high-throughput sequencing (Fig. 2 and Table 2). Among them, 15 putative viral genome sequences showed more than 90% sequence identity with known viruses and were classified as known viruses. The remaining 14 putative viral genome sequences showed less than 90% sequence identity with known viruses and were considered putative novel viruses (Table 2). All raw sequencing reads were deposited in the Sequence Read Archive (SRA) database: BioProject accession No. PRJNA1517747, BioSample accession numbers from SAMN62706723 to SAMN62706724, and SRA runs from SRR40363323 to SRR40363338.

Table 2
Information of putative virus detected in *Plasmodiophora brassicae* by high-throughput sequencing.

No.	Name	Accession number	Length	Sample	Best match	Identity (% aa)	Family	Type
1	Plasmodiophora brassicae RNA virus 1 (PbRV1)	PZ900343	11950	S2-2024	Maize suscal virus MZ270532	37.9	unclassified	+ssRNA
2	Plasmodiophora brassicae secovirus 1 (PbSV1)	PZ900342	9724	S2-2024	Triticum aestivum secovirus BK062998	69.9	Secoviridae	+ssRNA
3	Plasmodiophora brassicae dicistro-like virus 1 (PbDLV1)	PZ900344	8760	S2-2024	Tomato dicistro-like virus OQ983468	58.4	Dicistroviridae	+ssRNA
4	Turnip mosaic virus	-	8702	S2-2024	Turnip mosaic virus OP548622	98.8	Potyviridae	+ssRNA
5	Plasmodiophora brassicae fusagravirus 1 (PbFV1)	PZ900345	7148	S2-2024	Bremia lactucae associated fusagravirus 1 MN565680	50.3	Fusagraviridae	dsRNA
6	Brassica virus 1_Nap	-	6844	S2-2024	Brassica virus 1_Nap BK061753	92.4	Rhabdoviridae	-ssRNA
7	Brassica campestris chinensis coguvirus 1	-	6736	S2-2024	Brassica campestris chinensis coguvirus 1 NC_079045	99.5	Phenuiviridae	-ssRNA
8	Brassica virus 2_Jun	-	6421	S2-2024	Brassica virus 2_Jun BK061749	98.2	Rhabdoviridae	-ssRNA
9	Plasmodiophora brassicae bunya/phlebo-like virus 1 (PbBPLV1)	PZ900346	3692	S2-2024	Rhizoctonia solani bunya/phlebo-like virus 1 MK507779	64.7	Phenuiviridae	-ssRNA
10	Plasmodiophora brassicae yue-like virus 1 (PbYLV1)	PZ900347	3575	S2-2024	Goldenrod fern yue-like virus OX380389	55.8	Yueviridae	-ssRNA
11	Brassica rapa virus 1	-	3474	S2-2024	Brassica rapa virus 1 NC_079120	99.6	Rhabdoviridae	-ssRNA
12	Cucumber mosaic virus	-	3467	S2-2024	Cucumber mosaic virus HQ283392	96.8	Bromoviridae	+ssRNA
13	Qingdao RNA virus 2	-	3408	S2-2024	Qingdao RNA virus 2 LC726797	92.4	Narnaviridae	+ssRNA

No.	Name	Accession number	Length	Sample	Best match	Identity (% aa)	Family	Type
14	Turnip yellows virus	-	2860	S2-2024	Turnip yellows virus OQ377541	98.6	Solemoviridae	+ssRNA
15	Rhizoctonia solani mitovirus 88	-	2460	S2-2024	Rhizoctonia solani mitovirus 88 OM984461	93.7	Mitoviridae	+ssRNA
16	Brassica yellows virus	-	2117	S2-2024	Brassica yellows virus HQ388351	99.8	Solemoviridae	+ssRNA
17	Plasmodiophora brassicae partitivirus 1 (PbPV1)	PZ900348	2076	S2-2024	Trichoderma harzianum partitivirus 1 MG973751	83	Partitiviridae	dsRNA
18	Raphanus sativus cryptic virus 2	-	1653	S2-2024	Raphanus sativus cryptic virus 2 MZ437436	99.8	Partitiviridae	dsRNA
19	Brassica campestris chinensis cryptic virus 1	-	1568	S2-2024	Brassica campestris chinensis cryptic virus 1 MW026405	95.7	Partitiviridae	dsRNA
20	Plasmodiophora brassicae negative-stranded RNA virus 1 (PbNSRV1)	PZ900349	2237	S1-2023	Ananindeua virus KY013484	24.6	unclassified	-ssRNA
21	Beet western yellows virus associated RNA	-	2830	S1-2023	Beet western yellows virus associated RNA MT642437	97.1	Solemoviridae	+ssRNA
22	Plasmodiophora brassicae RNA virus 2 (PbRV2)	PZ900350	1320	S1-2023	Beihai narna-like virus 23 NC_032457	45	unclassified	+ssRNA
23	Brassica napus RNA virus 1	-	12280	S1-2023	Brassica napus RNA virus 1 NC_040586	99.5	Secoviridae	+ssRNA
24	Plasmodiophora brassicae circular virus 1 (PbcDV1)	PZ900351	1567	S1-2023	Callinectes sapidus associated circular virus NC_027785	27.9	unclassified	cDNA
25	Plasmodiophora brassicae negative-stranded RNA virus 2 (PbNSRV2)	PZ900352	2143	S1-2023	Isopteran arli-related virus OKIAV103 MT153546	30.3	unclassified	-ssRNA

No.	Name	Accession number	Length	Sample	Best match	Identity (% , aa)	Family	Type
26	Plasmodiophora brassicae RNA virus 3 (PbRV3)	PZ900353	3274	S1-2023	Nola virus MK896512	32.5	unclassified	+ssRNA
27	Plasmodiophora brassicae RNA virus 4 (PbRV4)	PZ900354	9826	S1-2023	Red clover RNA virus 1 MG596242	58.9	unclassified	+ssRNA
28	Plasmodiophora brassicae dsRNA virus 1 (PbDsRV1)	PZ900355	1232	S1-2023	Rice ragged stunt virus (isolate Thailand) Q86287	22.7	Spinareoviridae	dsRNA
29	Turnip yellows virus associated RNA	-	2836	S1-2023	Turnip yellows virus associated RNA MT642445	97.9	Solemoviridae	+ssRNA

The 29 putative viruses were classified into 13 families: *Mitoviridae*, *Narnaviridae*, *Partitiviridae*, *Phenuiviridae*, *Bromoviridae*, *Dicistroviridae*, *Potyviridae*, *Fusagraviridae*, *Rhabdoviridae*, *Secoviridae*, *Solemoviridae*, *Spinareoviridae*, and *Yueviridae*. Among them, 15 putative viruses were classified as +ssRNA type, eight as -ssRNA type, five as dsRNA type, and one as DNA type. Only six of the 29 putative viruses, including Cucumber mosaic virus, Turnip yellows virus, Brassica yellows virus, Plasmodiophora brassicae RNA virus 1, Plasmodiophora brassicae partitivirus 1, and Turnip mosaic virus, showed a detection frequency of 50% or higher. Most putative viruses were detected in one or two samples, indicating strong regional specificity in their distribution (Fig. 2).

3.2. Positive-sense single-stranded RNA viruses

A total of 15 putative viruses with +ssRNA genomes were identified in this study. They were classified into *Secoviridae* (n = 2), *Dicistroviridae* (n = 1), *Mitoviridae* (n = 1), *Narnaviridae* (n = 1), *Potyviridae* (n = 1), *Solemoviridae* (n = 4), and *Bromoviridae* (n = 1). The remaining four viruses belong to unclassified viruses (Fig. 2). A putative novel secovirus with a genome length of 9724 nt, named Plasmodiophora brassicae secovirus 1 (PbSV1), was identified in *P. brassicae* (Table 2). Genome annotation confirmed that this virus belongs to the family *Secoviridae* (order *Picornavirales*). Key viral functional domains, including RdRp, RNA helicase, and Peptidase_C3G, were predicted in the assembled viral sequences (Fig. 3). Phylogenetic analysis based on the amino acid sequence of the conserved RdRp region further supported that PbSV1 clusters within the *Secoviridae* clade and represents a distinct viral lineage divergent from Triticum aestivum secovirus (BK062998, 69.9% identity), which is currently classified within *Secoviridae*.

A putative novel dicistro-like virus with a genome length of 8760 nt, named Plasmodiophora brassicae dicistro-like virus 1 (PbDLV1), was identified in *P. brassicae* (Table 2). Genome annotation confirmed that this virus belongs to the family *Dicistroviridae*. Key viral functional domains, including RdRp, RNA helicase, and capsid protein, were predicted in the assembled viral sequences (Fig. 3). Phylogenetic analysis based on the amino acid sequence of the conserved RdRp region further supported that PbDLV1 clusters within the *Dicistroviridae* clade and represents a distinct viral lineage divergent from Tomato dicistro-like virus (OQ983468, 58.4% identity), which is currently classified within *Dicistroviridae*.

The remaining nine putative virus sequences, belonging to *Secoviridae* (n = 1), *Mitoviridae* (n = 1), *Narnaviridae* (n = 1), *Potyviridae* (n = 1), *Solemoviridae* (n = 4), and *Bromoviridae* (n = 1), showed greater than 90% sequence identity with Brassica napus RNA virus 1 (NC_040586), Qingdao RNA virus 2 (LC726797), Rhizoctonia solani mitovirus 88 (OM984461), Turnip mosaic virus (OP548622), Turnip yellows virus (OQ377541), Brassica yellows virus (HQ388351), Beet western

yellow virus associated RNA (MT642437), Turnip yellow virus associated RNA (MT642445), and Cucumber mosaic virus (HQ283392), respectively. These nine putative virus sequences were therefore identified as known viruses.

Four putative novel viruses, named *Plasmodiophora brassicae* RNA virus 1 (PbRV1, 11,950 nt), *Plasmodiophora brassicae* RNA virus 2 (PbRV2, 1320 nt), *Plasmodiophora brassicae* RNA virus 3 (PbRV3, 3274 nt), and *Plasmodiophora brassicae* RNA virus 4 (PbRV4, 9826 nt), were identified in *P. brassicae* (Table 2 and Fig. 4). The assembled genome sequences of putative viruses PbRV2 and PbRV3 were 1320 nt and 3274 nt in length, respectively. Due to the short-assembled sequences, sequence alignment and phylogenetic analysis could not further determine their taxonomic classification.

The assembled PbRV1 sequence showed 43.2% genome sequence identity with Maize suscal virus (MZ270532, 37.9% identity). The putative PbRV1 genome contains five ORFs and conserved domains including RdRp and mRNA-capping region (Fig. 4). The predicted RdRp amino acid sequence of PbRV1 showed 52.04% identity with that of Maize suscal virus. Phylogenetic analysis based on the RdRp amino acid sequence revealed that the putative virus PbRV1 is most closely related to Maize suscal virus, followed by viruses of the family *Lispiviridae*. Currently, the taxonomic status of Maize suscal virus remains unclear. Based on these structural features, PbRV1 was tentatively classified as an unclassified virus.

The assembled PbRV4 sequence showed 58.9% sequence identity with Red clover RNA virus 1 (MG596242, 7190 nt) (Table 2). The putative PbRV4 genome contains two ORFs and conserved domains including RdRp and RNA helicase (Fig. 4). Phylogenetic analysis based on the RdRp amino acid sequence revealed that PbRV4 is most closely related to Red clover RNA virus 1 and was tentatively classified as a Recovirid (*Benyviridae*-related viruses). Based on these structural features, PbRV1 was also tentatively classified as an unclassified virus.

3.3. Negative, single-stranded RNA viruses

Two putative novel viruses, named *Plasmodiophora brassicae* yue-like virus 1 (PbYLV1, 3575 nt) and *Plasmodiophora brassicae* bunya/phlebo-like virus 1 (PbBPLV1, 3692 nt), were identified in *P. brassicae* (Table 2). The assembled PbYLV1 genome sequence showed 55.8% identity with Goldenrod fern yue-like virus (OX380389, 3482 nt). The PbYLV1 genome contains one ORF, and no other conserved domains were detected (Fig. 5). Phylogenetic analysis based on the ORF amino acid sequence revealed that PbYLV1 is most closely related to Goldenrod fern yue-like virus and belongs to the family *Yueviridae*. Then, PbYLV1 was tentatively classified as *Yueviridae*.

The assembled PbBPLV1 genome sequence showed 64.7% sequence identity with *Rhizoctonia solani* bunya/phlebo-like virus 1 (MK507779, 7804 nt) (Table 2). Due to the incomplete assembly of the PbBPLV1 genome, only one ORF was identified, and no other conserved domains were detected. Phylogenetic analysis based on the ORF amino acid sequence revealed that PbBPLV1 is most closely related to *Rhizoctonia solani* bunya/phlebo-like virus 1 and belongs to the family *Phenuiviridae* (Fig. 5). Then, PbBPLV1 was tentatively classified as *Phenuiviridae*.

Another two putative novel viruses, named *Plasmodiophora brassicae* negative-stranded RNA virus 1 (PbNSRV1, 2237 nt) and *Plasmodiophora brassicae* negative-stranded RNA virus 2 (PbNSRV2, 2143 nt) with incomplete genome, were detected in *P. brassicae* (Table 2 and Fig. 5). They showed the highest sequence identity with Ananindeua virus (KY013484, 6924 nt) and Isopteran arli-related virus OKIAV103 (MT153546, 6724 nt) at 24.6% and 30.3%, respectively. Based on existing features, their taxonomic classification could not be determined.

Three putative virus sequences showed high identity (> 90%) with known viruses Brassica virus 1_Nap (BK061753, 6844 nt), Brassica virus 2_Jun (BK061749, 6421 nt), and Brassica rapa virus 1 (NC_079120, 3474 nt), which belong to *Rhabdoviridae* (n = 3) (Table 2). Another putative virus sequence showed high identity (> 90%) with known virus Brassica campestris chinensis coguvirus 1 (NC_079045, 6736 nt), which belongs to *Phenuiviridae*.

3.4. Double-stranded RNA viruses

Three putative novel double-stranded RNA viruses were detected in *P. brassicae* samples. The putative novel virus Plasmodiophora brassicae fusagravirus 1 (PbFV1, 7148 nt) showed 50.3% sequence identity with Bremia lactucae associated fusagravirus 1 (MN565680, 7210 nt) (Table 2). The PbFV1 genome contains three ORFs, with ORF2 encoding an RdRp conserved domain (Fig. 6). Phylogenetic analysis based on the RdRp amino acid sequence revealed that PbFV1 clusters within the *Fusagraviridae* clade and is most closely related to Bremia lactucae associated fusagravirus 1. Based on these features, PbFV1 was classified as a member of the family *Fusagraviridae*.

The putative novel virus Plasmodiophora brassicae partitivirus 1 (PbPV1, 2076 nt) genome contained only one segment, showing 83% sequence identity with Trichoderma harzianum partitivirus 1 (MG973751) segment 1 (Table 2). The PbPV1 genome segment 1 contains one ORF encoding an RdRp conserved domain (Fig. 6). Phylogenetic analysis based on the RdRp amino acid sequence revealed that PbPV1 is most closely related to Trichoderma harzianum partitivirus 1. Based on these features, PbPV1 was classified as a member of the family *Partitiviridae*.

The putative novel virus Plasmodiophora brassicae dsRNA virus 1 (PbDsRV1, 1232 nt) with incomplete genome, showing 22.7% sequence identity with Rice ragged stunt virus (isolate Thailand) (Q86287) (Table 2). Phylogenetic analysis based on the ORF amino acid sequence revealed a close relationship with the family *Spinareoviridae* (Fig. 6). PbDsRV1 was tentatively classified as a member of the family *Spinareoviridae*.

In addition, two putative virus sequences of 1653 nt and 1568 nt in length were detected and showed greater than 95% sequence identity with known Raphanus sativus cryptic virus 2 (MZ437436) and Brassica campestris chinensis cryptic virus 1 (MW026405), respectively (Table 2). These two putative viruses were therefore identified as known viruses.

3.5. Virus distribution and correlation analysis

The characteristic of virus distribution in 12 *P. brassicae* samples collected from 11 different regions in 2024 was analyzed. A total of 19 putative viruses were detected in these 12 samples, including seven novel viruses and 12 known viruses (Fig. 7A). Total of 15 putative viruses were detected in the KD sample, with PbFV1, PbDLV1, PbYLV1, Brassica campestris chinensis coguvirus 1, Raphanus sativus cryptic virus 2, and Brassica campestris chinensis cryptic virus 1 detected exclusively in the KD sample. Only two viruses named Cucumber mosaic virus and Brassica rapa virus 1 were detected in the Aba sample. Three putative viruses, named Cucumber mosaic virus, Turnip yellows virus, and Turnip mosaic virus, were detected in the XC sample. Approximately four to seven viruses were detected in the remaining samples.

Among them, the five most widely distributed viruses were Cucumber mosaic virus (100%), Turnip yellows virus (91.7%), Brassica yellows virus 1 (75%), PbPV1 (66.7%), and PbRV1 (66.7%). Cucumber mosaic virus and PbPV1 showed the highest relative abundance in the YB sample. Turnip yellows virus and Brassica yellows virus 1 showed the highest relative abundance in the Hdi sample from the Mianyang city. PbRV1 showed the highest relative abundance in the KD sample. Correlation analysis between viral relative abundance and environmental factors (disease incidence, longitude, latitude, and altitude) revealed that the relative abundance of PbPV1 and Brassica virus 1_Nap in the samples was significantly positively correlated with disease incidence (Fig. 7B, $p < 0.05$).

The Richness index, Shannon index, and Simpson index were calculated based on virus species and their relative abundance across samples to evaluate virus alpha diversity. Correlation analysis between *P. brassicae* virus alpha diversity indices and environmental factors (disease incidence, longitude, latitude, and altitude) revealed no significant correlations (Fig. 7C, $p > 0.05$), indicating that the viral diversity level in this study was not associated with the environmental factors examined at the sampling sites.

4. Discussion

This study represents the first investigation of the viral community in *P. brassicae* in Sichuan province and Guizhou province. Through high-throughput sequencing analysis, a total of 29 viruses were detected, significantly enriching our understanding of the viral diversity harbored by *P. brassicae*. Based on sequence alignment and phylogenetic analysis, 14 novel viruses were identified among the 29 viruses, belonging to different RNA viral taxonomic groups. The most abundant nucleic acid type among the viruses detected in *P. brassicae* was +ssRNA, involving seven viral families and 15 viruses, of which six were novel.

The family *Secoviridae* comprises +ssRNA viruses capable of infecting plants, with genomes organized as either monopartite or bipartite [24]. Similar to most *secoviruses*, the PbSV1 genome contains a single large ORF encoding RdRp, RNA helicase, and Peptidase_C3G, which are absent in *Triticum aestivum* secovirus [25]. *Triticum aestivum* secovirus was detected through high-throughput sequencing, and its complete genome structure has not been verified. *Triticum aestivum* secovirus is currently identified as a *waika-like* virus and has not been assigned to any of the six established genera within *Secoviridae*. Based on the similar genome structure and phylogenetic relationship between PbSV1 and *Triticum aestivum* secovirus, we suggest that PbSV1 may also belong to the *waika-like* virus group.

The putative novel virus PbDLV1 belongs to the family *Dicistroviridae* and showed the highest sequence identity (58.4%) with Tomato dicistro-like virus (TRaDLV). Similar to TRaDLV, the assembled PbDLV1 genome contains two ORFs. ORF1 encodes a capsid protein, while ORF2 encodes RdRp and RNA helicase. TRaDLV belongs to the *dicistro-like* virus genus and has only been detected in tomato root tissues, where it can induce yellow mosaic symptoms in leaves [26], suggesting a potential association with root microbe colonization. Based on the similar genome structure and phylogenetic relationship, we propose that PbDLV1 also belongs to the *dicistro-like* virus genus.

PbRV1, PbRV2, PbRV3, and PbRV4 were identified as unclassified viruses in this study. The PbRV1 genome showed the highest sequence identity (37.9%) with Maize suscal virus, which was detected through high-throughput sequencing in maize and whose genome structure has not been verified [27]. The PbRV1 genome shares similar characteristics with Maize suscal virus, containing five ORFs that encode an RdRp and an mRNA capping domain. Based on the predicted genome structure and phylogenetic analysis, PbRV1 was tentatively classified as an unclassified virus.

The assembled PbRV4 sequence showed 58.9% sequence identity with Red clover RNA virus 1 (RCRV1) (Table). RCRV1 was also detected through high-throughput sequencing in *Trifolium pratense* L. samples, with a total genome length of 7190 nt [28]. A recent study suggested that RCRV1, *Arceuthobium sichuanense* virus 3, and Goji berry chlorosis virus may represent single-component Reclovirids belonging to *Benyviridae*-related viruses [29, 30]. Viruses in the Reclovirid group can have genomes up to 10,000 nt, with genome structures and encoded proteins that differ from those of *Benyviridae*. In this study, the assembled PbRV4 sequence (9826 nt) is approximately 2.6 kb longer than the reported RCRV1 genome. Consistent with RCRV1, the assembled PbRV4 genome contains two ORFs, with the larger ORF2 encoding RdRp and RNA helicase. Although the PbRV4 genome sequence was not experimentally verified in this study, phylogenetic analysis showed that it is closely related to RCRV1, *Arceuthobium sichuanense* virus 3, and Goji berry chlorosis virus, and more distantly related to viruses of the family *Benyviridae*. Therefore, we suggest that PbRV4, like RCRV1, belongs to the Reclovirid group.

Yuevirids possess two segments of linear negative-sense RNA with a total length of 7.8–8.2 kb (large segment: 6.6–6.9 kb; small segment: 1.2–1.3 kb) [31]. In this study, only one PbYLV1 genome sequence was detected, with an assembled length of 3575 nt, showing the highest sequence identity (55.8%) with Goldenrod fern yue-like virus (partial, 3285 nt). This may be due to the low concentration of PbYLV1 in *P. brassicae* samples, as PbYLV1 was detected only in samples from the KD region (Fig. 6).

PbFV1 was classified as a member of the family *Fusagraviridae*, which showed 50.3% sequence identity with *Bremia lactucae* associated fusagravirus 1. *Bremia lactucae* associated fusagravirus 1 is a virus that specifically parasitizes pathogenic oomycetes [32]. Genome structure prediction revealed that PbFV1 shares a similar genome organization,

containing three ORFs with ORF2 encoding RdRp. Evolutionary analysis of this virus group suggests that their genome structures gradually undergo variations during the long-term evolutionary process of plant-pathogen interactions. This scenario is analogous to the relationship between cruciferous crops and *P. brassicae*.

The assembled PbPV1 genome sequence contained only dsRNA segment 1 (2076 nt), encoding an RdRp, and showed 83% sequence identity with *Trichoderma harzianum* partitivirus 1 (ThPV1) genome segment 1. ThPV1 is a member of the genus *Betapartitivirus* in the family *Partitiviridae*, with a genome consisting of two segments of similar size: segment 1 contains ORF1 encoding RdRp, and segment 2 contains ORF2 encoding a capsid protein [33]. The family *Partitiviridae* includes five genera with characteristic hosts for each genus: plants or fungi for genera *Alphapartitivirus* and *Betapartitivirus*; fungi for genus *Gammapartitivirus*; plants for genus *Deltapartitivirus*; and protozoa for genus *Cryspovirus* [34]. Phylogenetic analysis showed that PbPV1 clusters with multiple viruses of the genus *Betapartitivirus*; therefore, we propose that PbPV1 also belongs to the genus *Betapartitivirus* in the family *Partitiviridae*. Viruses of this genus are widely distributed in soil fungal communities, including *Trichoderma harzianum*, *Sclerotinia sclerotiorum*, *Rosellinia necatrix*, *Rhizoctonia solani*, and *Heterobasidion* [33, 35, 36, 37, 38]. The PbPV1 of *P. brassicae* may have been acquired through contact with plants or endophytic fungi during the process of root cell parasitism.

Among them, 21 of the 29 putative viruses identified in *P. brassicae* samples are plant viruses, while the remaining eight viruses may have hosts in soil fungi or animals. Among these plant viruses, several are typical pathogenic plant viruses, including Cucumber mosaic virus, Turnip mosaic virus, and Beet western yellows virus associated RNA (39,40,41). The detection of such a large number of plant viruses in *P. brassicae* samples may be related to its unique life cycle. During cruciferous crop growing seasons, the resting spores of *P. brassicae* germinate to form zoospores, which infect root hairs and enter root tissue cells, where they complete the entire sexual reproduction and proliferation process [42]. During this period, *P. brassicae* engages in extensive interactions with plant cells and root endophytic microorganisms, providing conditions favorable for virus infection events.

P. brassicae samples were collected from 16 different regions in Sichuan province. Only six viruses showed a detection frequency of 50% or higher, indicating that the distribution of most viruses exhibits regional specificity. For example, 15 viruses were detected in the KD (2024) sample, far exceeding the median number of virus species (median = 6), and six viruses were detected exclusively in this sample (Fig. 7A). In contrast, only two and three viruses were detected in samples from the ABa and XC regions, respectively. This may be related to the limited dispersal capacity of *P. brassicae* itself. Resting spores in soil have poor autonomous movement ability and typically rely on water flow, agricultural operations, and human transport for long-distance dispersal [1]. These unstable transmission pathways result in minimal exchange between pathogen populations across different regions. A similar case exists in *Polymyxa graminis*, an obligate intracellular plasmodiophorid protist that parasitizes root cells of cereal crops. Although non-pathogenic itself, *P. graminis* serves as a soil-borne vector for multiple *bymoviruses* and *furoviruses* [43, 44, 45].

Correlation analysis revealed that the relative abundance of PbPV1 and Brassica virus 1_Nap in *P. brassicae* samples was significantly positively correlated with disease incidence (Fig. 7B, $p < 0.05$), suggesting that the increase in gall numbers in field plots may facilitate the accumulation of these two viruses. This implies that *P. brassicae* may serve as a vector for PbPV1 and Brassica virus 1_Nap. However, the analysis of virus alpha diversity showed no significant correlation between *P. brassicae* virus diversity indices and environmental factors including disease incidence, longitude, latitude, and altitude (Fig. 7C). This result may be due to the relatively limited sample size, which did not fully capture potential environmental associations. Alternatively, it may suggest that the primary drivers of *P. brassicae* viral community composition and distribution are not these macro-environmental variables, but rather intrinsic factors such as the population genetic structure, physiological state, or microhabitat of the *P. brassicae* host. This study primarily obtained viral genome sequence information through high-throughput sequencing technology, representing a discovery and description phase. The specific transmission pathways, complete biological functions, and detailed interactions between these newly identified viruses and their *P. brassicae* hosts remain unknown and require further in-depth research.

5. Conclusion

The present study performed a meta-transcriptomic survey of viral communities in resting spores of *P. brassicae* collected from Sichuan and Guizhou Provinces, China. Twenty-nine putative viruses belonging to 13 viral families were identified, among which 14 are putative novel viral species. This constitutes the first comprehensive description of viruses associated with *P. brassicae*, substantially expanding the known viral diversity of obligate biotrophic protist. Greenhouse assays demonstrated clear virulence variation among geographically distinct *P. brassicae* populations. Correlation analyses showed that the relative abundances of PbPV1 and Brassica virus 1_Nap were significantly positively correlated with clubroot disease incidence. Most viruses exhibited notable geographic endemism, reflecting local pathogen population characteristics. This study resolves a key knowledge gap in the clubroot pathosystem and provides a foundational dataset for future research on virus-*P. brassicae* interactions. Further functional investigations are required to elucidate viral transmission and their modulatory effects on pathogen fitness, which will deepen our understanding of clubroot epidemiology and may inform novel disease management approaches.

Abbreviations

RdRp RNA dependent RNA polymerase

NCBI National Center for Biotechnology Information

CD Conserved Domain

ORF Open Reading Frame

TPM Transcripts per million

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This work was supported by the National Natural Science Foundation of China (32300016), Sichuan Provincial Natural Science Foundation (2026NSFSC1044), the Talent Research Fund (NKYRCZX2024018), the Science and Technology Project (1 + 9KJGG006) and the Scientific and Technological Achievement Transformation Project (2026ZSSFGH23) of Sichuan Academy of Agricultural Sciences, and Sichuan Rape Innovation Team Project of the Modern Agricultural Industry Technology System (SCCXTD-2026-3).

Author Contribution

Yue Deng wrote the main manuscript text, Data analysis was completed by Yue Deng, Yunhui He, Xiuying Qian, and Kang Zhou. Yaoying Yu, Xiaoqin Huang, Xiaoxiang Yang, Peisong, and Yunxi Xian jointly completed the collection, processing, and preservation of samples. Funding sources were managed by Yue Deng, Yong Liu, and Lei Zhang. Yue Deng, Yong Liu, Kang Zhou, and Lei Zhang reviewed and revised the manuscript.

Acknowledgements

The authors thank the Rapeseed Research Institute, Guizhou Academy of Agricultural Sciences for providing the clubroot samples.

Availability of data and materials

The datasets generated during the current study are available in the National Center for Biotechnology Information (NCBI) database (accession numbers are PRJNA1517747).

References

1. Javed MA, Schwelm A, Zamani-Noor N, Salih R, Silvestre Vañó M, Wu J, et al. The clubroot pathogen *Plasmodiophora brassicae*: A profile update. *Mol Plant Pathol*. 2023;24:89–106. 10.1111/mpp.13283.
2. Chai AL, Xie XW, Shi YX, Li BJ. Research status of clubroot (*Plasmodiophora brassicae*) on cruciferous crops in China. *Can J Plant Pathol*. 2014;36(sup1):142–53. 10.1080/07060661.2013.868829.
3. Dixon GR. The occurrence and economic impact of *Plasmodiophora brassicae* and clubroot disease. *J Plant Growth Regul*. 2009;28:194–202. 10.1007/s00344-009-9090-y.
4. Brader G, Compant S, Vescio K, Mitter B, Trognitz F, Ma LJ, Sessitsch A. Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annu Rev Phytopathol*. 2017;55:61–83. 10.1146/annurev-phyto-080516-035641.
5. Zhou K, Deng Y, Zhu C, Yang L, Zhang J, Chen W, Suzuki N, Li G, Wu M. Core virome shapes adaptation of a phytopathogenic fungus to climate and cropping patterns. *ISME J*. 2026;20(1):wrag001. 10.1093/ismejo/wrag001.
6. Prosdocimi F, Cortines JR, José MV, Farias ST. Decoding viruses: An alternative perspective on their history, origins and role in nature. *BioSystems*. 2023;231:104960. 10.1016/j.biosystems.2023.104960.
7. Dolja VV, Koonin EV. Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res*. 2018;244:36–52. 10.1016/j.virusres.2017.10.020.
8. Watanabe T, Suzuki N, Tomonaga K, et al. Neo-virology: the raison d'être of viruses. *Virus Res*. 2019;274:197751. 10.1016/j.virusres.2019.197751.
9. Mojica KDA, Huisman J, Wilhelm SW, Brussaard CPD. Latitudinal variation in virus-induced mortality of phytoplankton across the North Atlantic Ocean. *ISME J*. 2016;10:500–13. 10.1038/ismej.2015.130.

10. Bachy C, Yung CCM, Needham DM, Gazitúa MC, Roux S, Limardo AJ, et al. Viruses infecting a warm water picoeukaryote shed light on spatial co-occurrence dynamics of marine viruses and their hosts. *ISME J*. 2021;15:3129–47. 10.1038/s41396-021-00989-9.
11. Philippe C, Bernard LS, Didier R. Giant Viruses of Amoeboe: A Journey Through Innovative Research and Paradigm Changes. *Annual Rev Virol*. 2017;4(1):61–85. 10.1146/annurev-virology-101416-041816.
12. Kuhne T. Soil-borne viruses affecting cereals: known for long but still a threat. *Virus Res*. 2009;141:174–83. 10.1016/j.virusres.2008.05.019.
13. Strelkov SE, Tewari JP, Smith-Degenhardt E. (2006). Characterization of *Plasmodiophora brassicae* populations from Alberta, Canada. *Canadian Journal of Plant Pathology*. 2006; 28: 467–474; doi.org/10.1080/07060660609507321
14. Kuginuki Y, Yoshikawa H, Hirai M. Variation in virulence of *Plasmodiophora brassicae* in Japan tested with clubroot-resistant cultivars of chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Eur J Plant Pathol*. 1999;105:327–32. doi.org/10.1023/A:1008705413127.
15. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20. 10.1093/bioinformatics/btu170.
16. Karlsson FH, Fåk F, Nookaew I, et al. Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nat Commun*. 2012;3:1245. 10.1038/ncomms2266.
17. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. MEGAHIT v1.0: A Fast and Scalable Metagenome Assembler driven by Advanced Methodologies and Community Practices. *Methods*. 2016. 10.1016/j.ymeth.2016.02.020.
18. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics*. 2010;25(14):1754–60. 10.1093/bioinformatics/btp324.
19. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12:59–60. 10.1038/nmeth.3176.
20. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14:417–9. 10.1038/nmeth.4197.
21. Wille M, Harvey E, Shi M, Gonzalez-Acuna D, Holmes EC, Hurt AC. Sustained RNA virome diversity in Antarctic penguins and their ticks. *ISME J*. 2020;14:1768–82. 10.1038/s41396-020-0643-1.
22. Wille M, Shi M, Klaassen M, Hurt AC, Holmes EC. Virome heterogeneity and connectivity in waterfowl and shorebird communities. *ISME J*. 2019;13:2603–16. 10.1038/s41396-019-0458-0.
23. Oksanen J, Guillaume BF, Friendly M, Kindt R, Legendre P, McGlinn D, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, et al. Vegan: Community Ecology Package. *R Package*. 2022. 10.32614/CRAN.package.vegan. Version 2.6-2.
24. Fuchs M, Hily JM, Petrzik K, Sanfaçon H, Thompson JR, van der Vlugt R, Wetzel T, ICTV Report Consortium. *J Gen Virol*. 2022;103(12):001807. 10.1099/jgv.0.001807.
25. Sidharthan VK, Reddy V, Kiran G, et al. Probing of plant transcriptomes reveals the hidden genetic diversity of the family *Secoviridae*. *Arch Virol*. 2024;169:150. 10.1007/s00705-024-06076-6.
26. Cao X, Wang Z, Pang J, et al. Identification of a novel dicistro-like virus associated with the roots of tomato plants. *Arch Virol*. 2023;168:214. 10.1007/s00705-023-05843-1.
27. Li Q, Yang L, Zhu T, Yv X, Zhang B, Li H, Hao J, Zhang L, Ji P, Dong J. Detection and discovery of plant viruses in *Disporopsis* through high-throughput sequencing. *Front Microbiol*. 2024;15:1434554. 10.3389/fmicb.2024.1434554.
28. Koloniuk I, Fránová J. Complete nucleotide sequence and genome organization of red clover associated virus 1 (RCaV1), a putative member of the genus *Waikavirus* (family *Secoviridae*, order *Picornavirales*). *Archives of Virology*. 2018;163:3447–3449; 10.1007/s00705-018-4005-5

29. Solovyev AG, Morozov SY. Uncovering Plant Virus Species Forming Novel Provisional Taxonomic Units Related to the Family *Benyviridae*. *Viruses*. 2022;14(12):2680. 10.3390/v14122680.
30. Morozov SY, Lezzhov AA, Solovyev AG. Predicted Membrane-Associated Domains in Proteins Encoded by Novel Monopartite Plant RNA Viruses Related to Members of the Family *Benyviridae*. *Int J Mol Sci*. 2023;24:12161. 10.3390/ijms241512161.
31. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ. Redefining the invertebrate RNA virosphere. *Nature*. 2016;540(7634):539–43. 10.1038/nature20167.
32. Forgia M, Daghino S, Chiapello M, Ciuffo M, Turina M. New clades of viruses infecting the obligatory biotroph *Bremia lactucae* representing distinct evolutionary trajectory for viruses infecting oomycetes. *Virus Evol*. 2024;10(1):veae003. 10.1093/ve/veae003.
33. Chun J, Yang H-E, Kim D-H. Identification of a novel Partitivirus of *Trichoderma harzianum* NFCF319 and evidence for the related antifungal activity. *Front Plant Sci*. 2018;9:1699. 10.3389/fpls.2018.01699.
34. Vainio EJ, Chiba S, Ghabrial SA, Maiss E, Roossinck M, Sabanadzovic S, Suzuki N, Xie J, Nibert M, ICTV Report Consortium. ICTV Virus Taxonomy Profile: *Partitiviridae*. *J Gen Virol*. 2018;99:17–8. 10.1099/jgv.0.000985.
35. Xiao X, Cheng J, Tang J, Fu Y, Jiang D, et al. A novel partitivirus that confers hypovirulence on plant pathogenic fungi. *J Virol*. 2014;88:10120–33. 10.1128/jvi.01036-14.
36. Chiba S, Lin YH, Kondo H, Kanematsu S, Suzuki N. Effects of defective interfering RNA on symptom induction by, and replication of, a novel partitivirus from a phytopathogenic fungus, *Rosellinia necatrix*. *J Virol*. 2013;87:2330–41. 10.1128/jvi.02835-12.
37. Chu H, Shi R, Wu C, et al. Complete genome sequence of a novel partitivirus identified in *Rhizoctonia solani* AG-4 HG III isolate Rs31. *Arch Virol*. 2025;170:251. 10.1007/s00705-025-06443-x.
38. Vainio EJ, Hantula J. Taxonomy, biogeography and importance of *Heterobasidion* viruses. *Virus Res*. 2016;219:2–10. 10.1016/j.virusres.2015.10.014.
39. Joshi M, Narute TK, Sarnobat D. The Cucumber Mosaic Virus: a review. *Int J Plant Soil Sci*. 2023;35(24):253–60. 10.9734/ijpss/2023/v35i244316.
40. Cuesta R, Yuste-Calvo C, Gil-Cartón D, Sánchez F, Ponz F, Valle M. Structure of Turnip mosaic virus and its viral-like particles. *Sci Rep*. 2019;9:15396. 10.1038/s41598-019-51823-4.
41. Passmore BK, Sanger M, Chin LS, Falk BW, Bruening G. Beet western yellows virus-associated RNA: an independently replicating RNA that stimulates virus accumulation. *Proceedings of the National Academy of Sciences USA*. 1993;90(21):10168–10172; 10.1073/pnas.90.21.10168
42. Liu L, Qin L, Zhou Z, Hendriks WGHM, Liu S, Wei Y. Refining the life cycle of *Plasmodiophora brassicae*. *Phytopathology*. 2020;110(10):1704–12. 10.1094/PHYTO-02-20-0029-R.
43. Ward E, Fenn MGE, Henry CM. A rapid method for direct detection of *Polymyxa* DNA in soil. *Plant Pathol*. 2004;53:485–90. 10.1111/j.1365-3059.2004.01017.x.
44. Ratti C, Budge G, Ward L, Clover G, Rubies AC, Henry C. Detection and relative quantitation of Soil-borne cereal mosaic virus (SBCMV) and *Polymyxa graminis* in winter wheat using real-time PCR (Taqman). *J Virol Methods*. 2004;122:95–103. 10.1016/j.jviromet.2004.08.013.
45. Kanyuka K, Ward E, Adams MJ. *Polymyxa graminis* and the cereal viruses it transmits: a research challenge. *Mol Plant Pathol*. 2003;4:393–406. 10.1046/j.1364-3703.2003.00177.x.

Figures

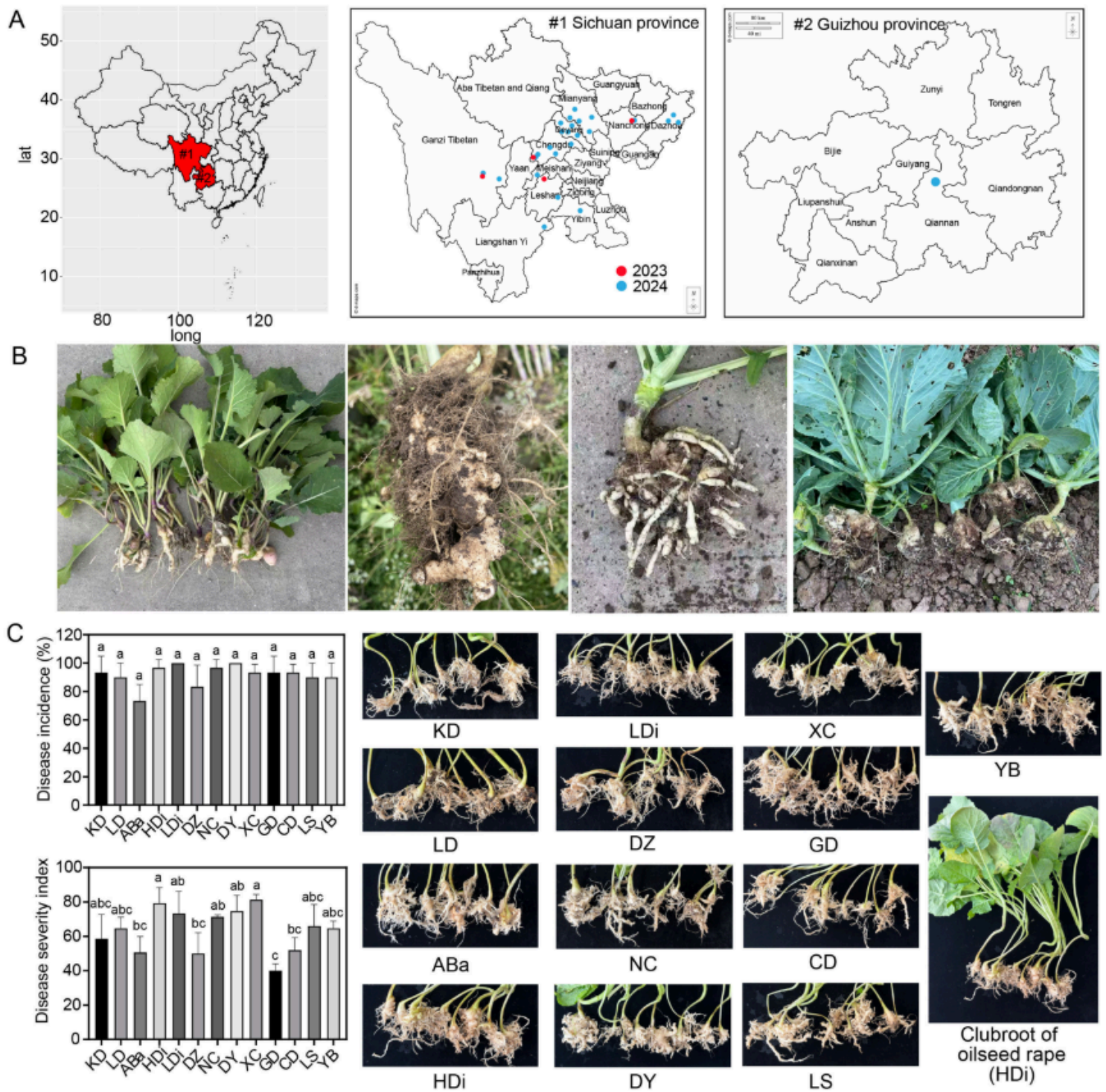


Figure 1

Sample collection and pathogenicity analysis of *Plasmodiophora brassicae*. (A) and (B) Sampling location information and clubroot symptoms of diseased samples. (C) Pathogenicity assay of collected clubroot samples under greenhouse conditions.

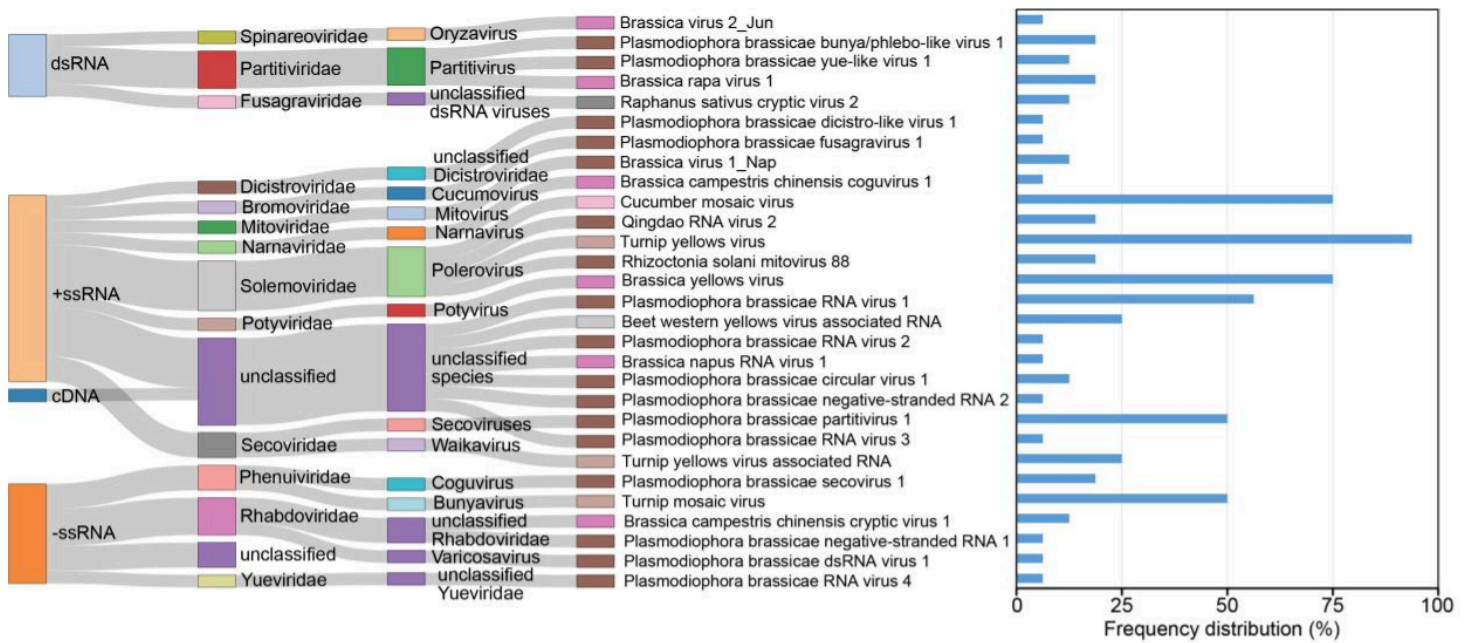


Figure 2

Sankey diagram of putative virus species detected in *P. brassicae*.

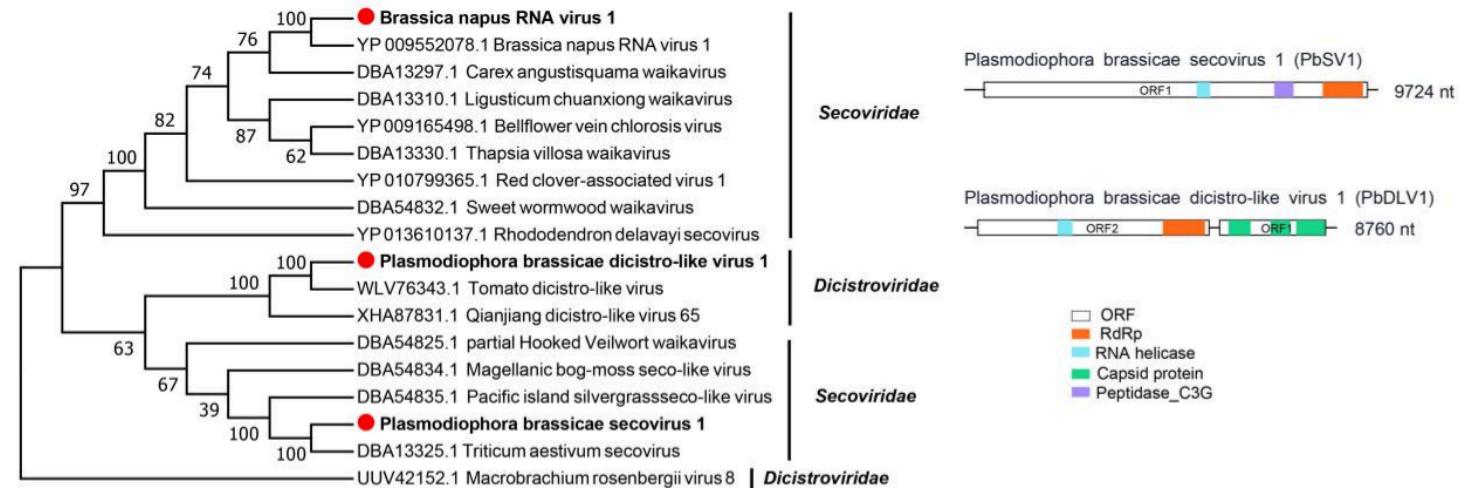


Figure 3

Genomic organization and phylogenetic analysis of putative novel viruses belonging to the families *Secoviridae* and *Dicistroviridae*. The phylogenetic tree was generated using the RdRp aa sequences and the neighbor-joining method with 1000 bootstraps. The bootstrap values are indicated adjacent to the nodes.

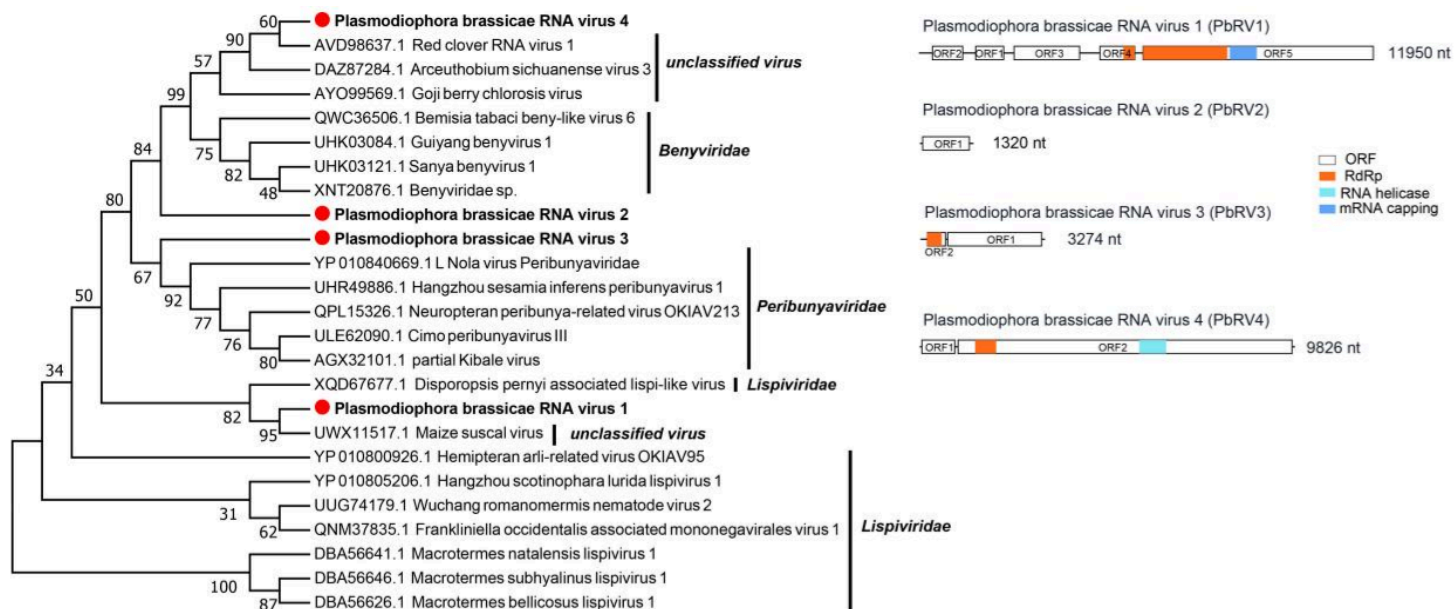


Figure 4

Genomic organization and phylogenetic analysis of putative novel viruses belonging to unclassified virus. Phylogenetic trees were constructed using the neighbor-joining method with 1000 bootstraps based on RdRp aa sequences for all viruses, except for PbRV2, for which the ORF1 aa sequence was used. The bootstrap values are indicated adjacent to the nodes.

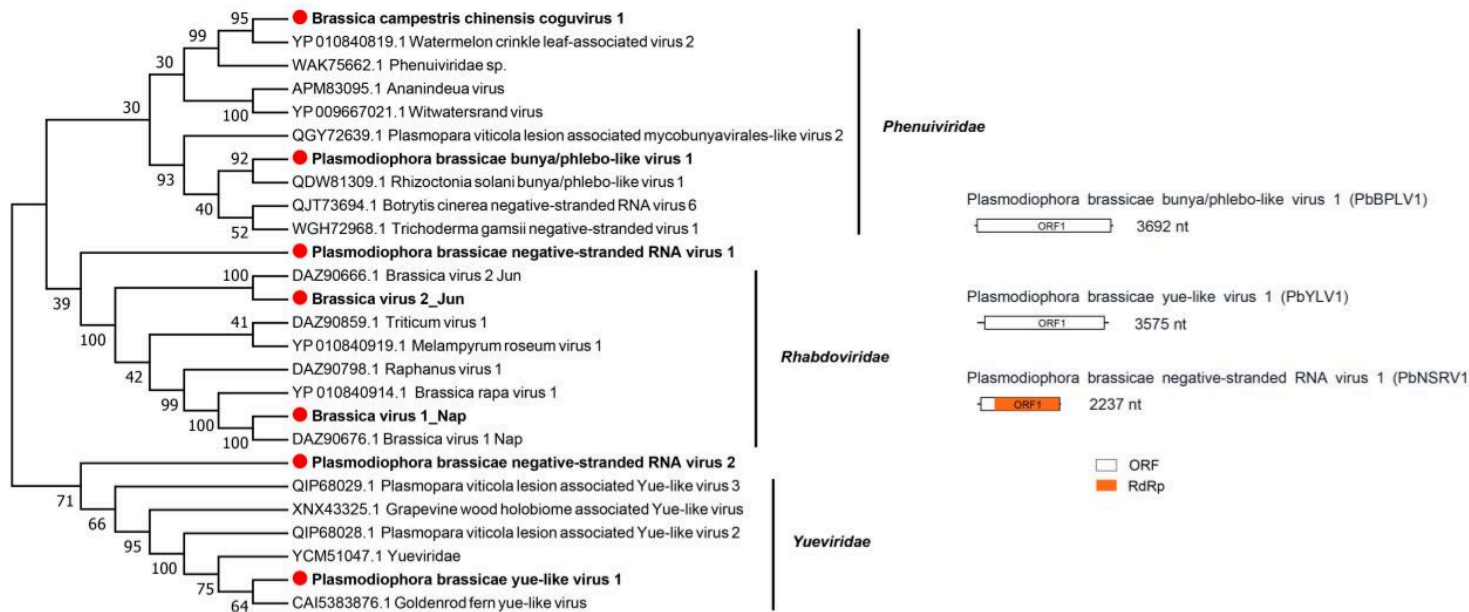


Figure 5

Genomic organization and phylogenetic analysis of putative novel negative single-stranded RNA viruses. The phylogenetic tree was generated using the ORF1 aa sequences and the neighbor-joining method with 1000 bootstraps. The bootstrap values are indicated adjacent to the nodes.



Figure 6

Genomic organization and phylogenetic analysis of putative novel double-stranded RNA viruses. The phylogenetic tree was generated using the RdRp aa sequences and the neighbor-joining method with 1000 bootstraps. The bootstrap values are indicated adjacent to the nodes.

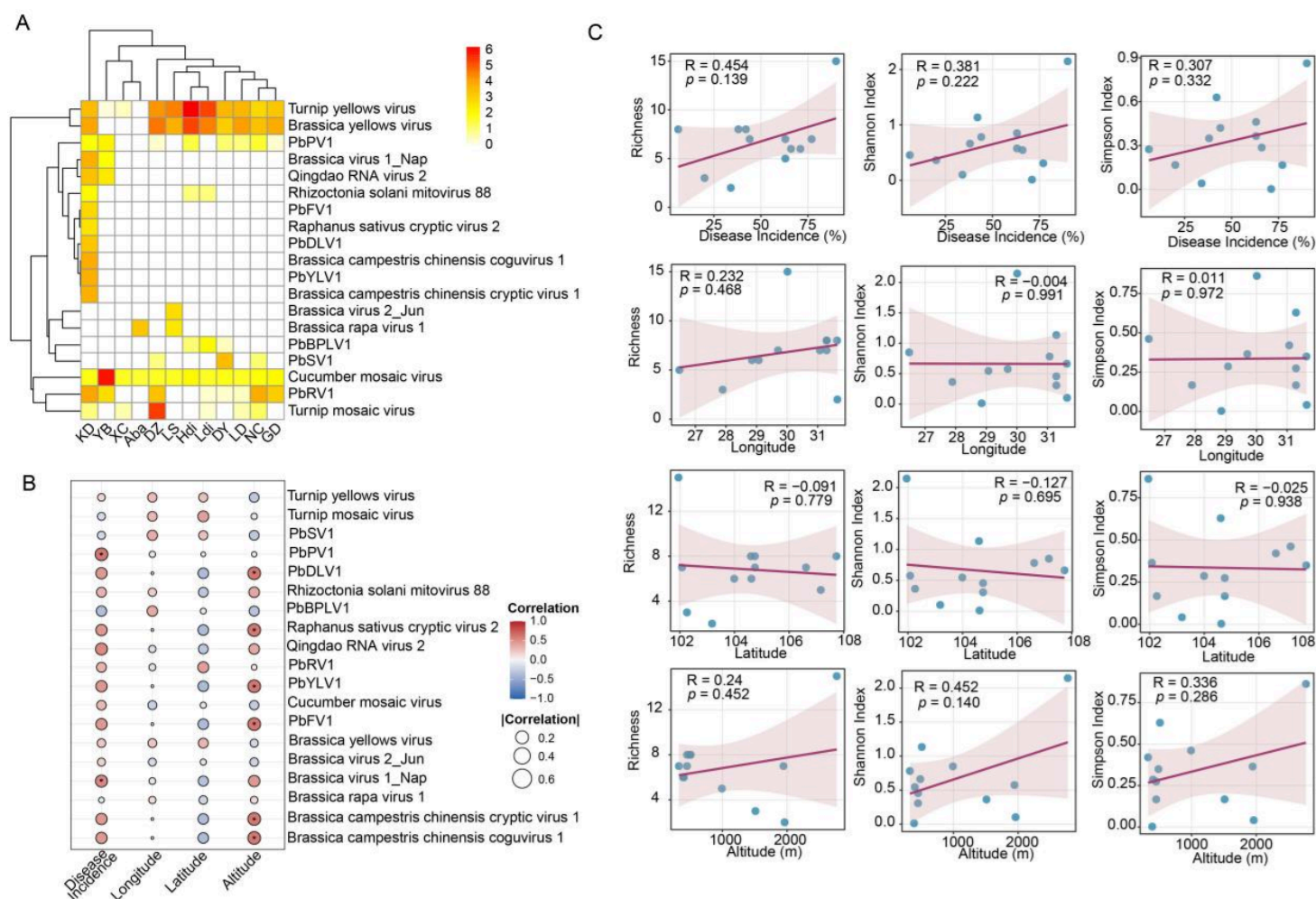


Figure 7

Viral distribution in *P. brassicae* samples from different regions and its correlation analysis with environmental factors.

(A) Heatmap showing the relative abundance of 19 putative viruses across 12 *P. brassicae* samples. (B) Correlation analysis between the relative abundance of 19 putative viruses and environmental factors including disease incidence, longitude, latitude, and altitude, at sampling sites. (C) Correlation analysis between the viruses alpha diversity index (Richness, Shannon index, and Simpson index) and environmental factors. * indicate significant correlation at $p < 0.05$.