

Article

Genomic Characterisation of an Isolate of *Brassica Yellow Virus* Associated with Brassica Weed in Tasmania

Muhammad Umar¹, Tahir Farooq² , Robert S. Tegg¹, Tamarasan Thangavel^{1,3} and Calum R. Wilson^{1,*} 

¹ New Town Research Laboratories, Tasmanian Institute of Agriculture, University of Tasmania, 13 St. Johns Avenue, New Town, TAS 7008, Australia; m.umar@utas.edu.au (M.U.); robert.tegg@utas.edu.au (R.S.T.); tamil.thangavel@daf.qld.gov.au (T.T.)

² Guangdong Provincial Key Laboratory of High Technology for Plant Protection, Plant Protection Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China; tfarooq@gdppri.com

³ Department of Agriculture and Fisheries (Queensland), Bundaberg Research Facility, 49 Ashfield Road, Bundaberg, QLD 4670, Australia

* Correspondence: calum.wilson@utas.edu.au; Tel.: +613-6226-6381

Abstract: *Brassica yellow virus* (BrYV), a tentative species in the genus *Polerovirus*, of the *Solemoviridae* family, is a phloem-restricted and aphid-transmitted virus with at least three genotypes (A, B, and C). It has been found across mainland China, South Korea, and Japan. BrYV was previously undescribed in Tasmania, and its genetic variability in the state remains unknown. Here, we describe a near-complete genome sequence of BrYV (genotype A) isolated from *Raphanus raphanistrum* in Tasmania using next-generation sequencing and sanger sequencing of RT-PCR products. BrYV-Tas (GenBank Accession no. OM469309) possesses a genome of 5516 nucleotides (nt) and shares higher sequence identity (about 90%) with other BrYV isolates. Phylogenetic analyses showed variability in the clustering patterns of the individual genes of BrYV-Tas. Recombination analysis revealed beginning and ending breakpoints at nucleotide positions 1922 to 5234 nt, with the BrYV isolate LC428359 and BrYV isolate KY310572 identified as major and minor parents, respectively. Results of the evolutionary analysis showed that the majority of the codons for each gene are evolving under purifying selection, though a few codons were also detected to have positive selection pressure. Taken together, our findings will facilitate an understanding of the evolutionary dynamics and genetic diversity of BrYV.

Keywords: *Brassica yellow virus*; genome characterisation; recombination; *Polerovirus*; phylogenetic analysis



Citation: Umar, M.; Farooq, T.; Tegg, R.S.; Thangavel, T.; Wilson, C.R. Genomic Characterisation of an Isolate of *Brassica Yellow Virus* Associated with Brassica Weed in Tasmania. *Plants* **2022**, *11*, 884. <https://doi.org/10.3390/plants11070884>

Academic Editor: Abdelali Hannoufa

Received: 31 January 2022

Accepted: 22 March 2022

Published: 25 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The genus *Polerovirus* (family *Solemoviridae*) encompasses ~32 species (including both tentative and approved) of genetically diverse, and economically important plant viruses [1]. *Brassica yellow virus* (BrYV) is a tentative polerovirus that was first reported infecting brassica spp. in China in 2011 [2]. Since then, it has been identified in other countries, including South Korea in 2015 [3], Japan in 2019 [4], and Australia in 2021 [5]. Similar to other poleroviruses, BrYV is a phloem-limited virus that is transmitted by specific aphid species in a persistent circulative manner. Several cruciferous plant species may be naturally infected by BrYV, including oilseed rape (*Brassica napus*), Chinese cabbage (*B. chinensis*), cabbage (*B. oleracea*), cauliflower (*B. oleracea* var. *botrytis*), rutabaga (*B. napobrassica*), leaf mustard (*B. juncea*), radish (*Raphanus sativus* var. *oleifera*), and white glabrous mustard (*B. alboglabara*). BrYV infection in crucifer plants induces mottling, yellowing, and leaf malformation symptoms [2,3].

The full-length genome of BrYV consists of a 5.6–5.7 kb long positive-sense single-stranded RNA (+ssRNA) which contains six open reading frames (ORFs), a short untranslated region (UTR) at 5', an intergenic non-coding region (NCR) spanning over ORF2 and ORF3, and a UTR at the 3' end that lacks a poly (A) tail or tRNA-like structure. The

genomic RNA is packed into a spherical-shaped virion of 25–30 nm diameter. Three unique genotypes, BrYV-A, BrYV-B, and BrYV-C, with genome sizes of 5666, 5666, and 5678 nt, respectively, have been described [2,6]. Additionally, a recombinant strain (BrYV-CS) has been reported in Chinese cabbage in South Korea [3]. Like typical poleroviruses, ORFs 0–2 are translated from genomic RNA, while ORFs 3–5 are translated from sub-genomic RNA (Figure 1). ORF0 expresses the P0 protein that functions as a determinant of host range and symptoms and acts as a silencing suppressor. The fusion of ORF1 and ORF2 encodes the P1–2 fusion protein or replication-associated protein (RepA). ORF3 encodes the major P3 or coat protein (CP) which sits in-frame with ORF5. The P3–5 overlapped translation results in the expression of a read-through protein (P5) which resides on the external surface of the mature virion and is associated with the long-distance movement and aphid transmission of the virus [7]. The ORF4 encodes the P4 protein, which plays an important role in virus systemic movement within infected plants (Figure 1).

BrYV (5.6–5.7 nt)

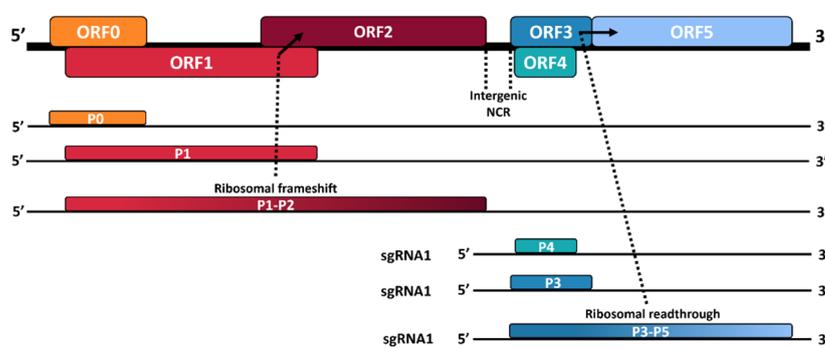


Figure 1. Representation of genomic organisation of BrYV. The ribosomal frameshift and read-through strategies are indicated by arrows; coloured boxes indicate the positions and types of different ORFs and translated proteins.

A recent study in Australia demonstrated that the highest genomic variability in BrYV, and its closely relative *Turnip yellows virus* (TuYV), occurred in ORF5 [5]. Based on phylogenetic and recombination analyses, this study proposed that BrYV and TuYV could be regarded as variants of the same species. However, a number of other studies have examined the genetic variation across other ORFs which differentiate BrYV from TuYV [2,3,6].

RNA recombination predominantly drives genetic diversity and plays a crucial role in the emergence, evolution, and epidemiology of the plant-infecting RNA viruses. It appears to be the main mechanism behind the formation of new and more pathogenic viral strains/species [8,9]. The genera *Polerovirus* and *Enamovirus* of the *Solemoviridae* family (previously in the family *Luteoviridae*) [10] are known to be commonly associated with recombination events [9]. Based on the phylogenetic analysis, several studies have demonstrated that self-recombination among poleroviruses or recombination with other luteoviruses has led to the observed evolution of poleroviruses [11–13].

Recently, complete genome sequencing of different plant viruses has expanded our understanding of their molecular characterisation, genetic diversity, and evolutionary relationships [14–16]. Here, we took advantage of next-generation sequencing (NGS) to obtain a near-complete genome sequence of a BrYV isolate infecting *R. raphanistrum* in Tasmania. We employed RT-PCR and multiplex RT-PCR to fill sequence gaps and confirm the BrYV genotype, respectively. Furthermore, we performed phylogenetic and recombination detection analyses to improve our understanding of the existing genetic variability and evolutionary relatedness between the newly described isolate and other globally reported BrYV isolates. The findings of our study will provide valuable information regarding

the genomic diversity and recombinational dynamics of BrYV to better understand and manage plant virus diseases.

2. Results

2.1. NGS Analysis

In total, 22,351,000 raw reads were obtained with Illumina sequencing for the tested sample. After quality trimming using BBDuK, these numbers were reduced to 22,228,228. De novo assembly generated sixteen contigs (ranged 3995–121 bp with N50 = 511 bp), produced by 670 viral reads. A BLASTn/BLASTx search of the GenBank database found a viral contig of ~3995 bp in sizes corresponding to the BrYV isolate “TO3” (Accession no. LC428360), sharing 96.68% nucleotide similarity.

2.2. Validation of NGS Results and Amplification of Missing Genome Sequence

Since the tested sample was a composite from five different plants, this necessitated further testing of individual leaves to determine the identification of the virus source. BrYV infection was confirmed in one *R. raphanistrum* sample by multiplex RT-PCR. Multiplex RT-PCR gave a amplicon of 277 bp which also revealed that the BrYV-Tas isolate belongs to genotype-A. To amplify the missing parts of the BrYV-Tas genome (i.e., 5'UTR, P0, and part of P1-P2 towards 5' end), RT-PCR assays were conducted with primers BrY-Ntab001F (+)/TuYVOrf0R and primer BrY-Ntab724F(+)/BrY-Ntab2064R(-) (Table 1), using cycling conditions, as previously described [17] with products sequenced. Consequently, the entire genome sequence of the BrYV Tasmanian isolate was determined, except for 138 bp at the termini of the 3'UTR (GenBank Accession no. OM469309).

Table 1. Primers used for RT-PCR, multiplex RT-PCR, and sequencing.

Primer Name	Sequence 5' to 3'	Target	Position	Product Size (bp)	PCR Conditions	References
BrY-Ntab001F (+)	ACAAAAGAAACC AGGAGGAAA	BrYV ORF0	1	780	1 × 95 °C (15 min), 35 × [94 °C (30 s), 55 °C (30 s), 72 °C (1 min)], 1 × 72 °C (10 min)	[17,18]
TuYVOrf0R	TCATACAAACATTT CGGTGTAGAC		781			
BrY-Ntab724F (+)	TCTCACTCCTGAA GAAATCC	BrYV ORF1-ORF2	724	1349	1 × 95 °C (15 min), 35 × [94 °C (30 s), 55 °C (30 s), 72 °C (1 min)], 1 × 72 °C (10 min)	[17]
BrY-Ntab2064R (-)	TGAATCACACGCTC CCTCTCAG		2073			
BrYA484F	TACTTGGACTAGAG ATGCTGAAAG	BrYV-ORF0 (Genotype A)	484	277	1 × 95 °C (10 min), 32 × [94 °C (30 s), 62 °C (60 s), 72 °C (1 min)], 1 × 72 °C (10 min)	[19]
BrYB88F	CCTCCACCCAAAAC AAGTAT	BrYV-ORF0 (Genotype B)	88	673		
BrYC257F	CGAGTTTCCGTACTT GTTG	BrYV-ORF0 (Genotype C)	257	504		
BrY761R	AGACCGAAGAGCT GAAAAGG	BrYV-ORF0 (Reverse)	761	-		

2.3. Multiple Sequence Alignment and Phylogenetic Analysis

The genomic sequence obtained for BrYV-Tas was 5516 nt in length and had a typical BrYV genomic organisation, with six ORFs (ORF0–ORF5), a 31 nt 5'-UTR, and an incomplete 47 nt at 3'-UTR. ORF0 spans nt 32 to 781 and is predicted to encode the P0 suppressor. ORF1 stretches between nt 174 to nt 1967 and encodes the P1 protein. ORF2 (nt 1853–3268) is assumed to be translated into a P1-P2 fusion protein (RdRp) via a –1 ribosomal frameshift. GGGAAAG at nt 1617 to 1623 is the shifty sequence. ORF3 starts at nt 3460 to 4068, encoding a putative coat protein (CP). ORF4 starts at nt 3491 to 4018 and codes for a putative P4 protein by a leaky scanning. ORF5 (from nt 4255 to 5469) is translated into the P3-P5 fusion protein via a read-through strategy, which is thought to play a key role in virus transmission (Figure 1).

Sequence homology and phylogenetic analysis based on the full genome indicated that the BrYV-Tas shares higher sequence identities (94.4%) with two Japanese BrYV isolates LC428360 and LC428361 and they cluster together in the same clade (II) (Figure 2A)

(Table S1). These Japanese isolates were found from *B. rapa* and *B. napus* hosts, respectively. The lowest homology percentage (81.3%) was observed for an isolate from China (MK057527) infecting tobacco. BrYV-Tas shared only 78.2% of nucleotide similarity with the closest TuYV isolate genomic sequence available on GenBank (NC_003743) (Table S1).

Subsequent alignments indicated that the P0, P1, P1-P2, P3, P3-P5, P4, and P5 genes of BrYV-Tas shared sequence identities of 90.9–98.1% for P0 (Table S2), 87.1–96.6% for P1 (Table S3), 89.3–92.1% for P1-P2 (Table S4), 92.6–98.9% for P3 (Table S5), 75.3–98.5% for P3-P5 (Table S6), 92.0–99.1% for P4 (Table S7), 67.1–98.8% for P5 (Table S8) with the corresponding genes of other BrYV isolates examined. Notably, phylogenetic analysis revealed that the clustering pattern of the individual BrYV-Tas genes, mentioned above, was variable (Figure 2). The percentage of nucleotide sequence identities between BrYV-Tas and TuYV (NC_003743) was 88.1% for P0 (Table S2), 86.6% for P1 (Table S3), 88.9% for P1-P2 (Table S4), 95.2% for P3 (Table S5), 67.0% for P3-P5 (Table S6), 95.5% for P4 (Table S7), and 55.2% for P5 (Table S8). Meanwhile, the amino acid percentage identities between BrYV-Tas and TuYV (NC_003743) was 82.7% for P0 (Table S9), 85.0% for P1 (Table S10), 88.9% for P1-P2 (Table S11), 95.0% for P3 (Table S12), 60.9% for P3-P5 (Table S13), 90.9% for P4 (Table S14), and 43.2% for P5 (Table S15).

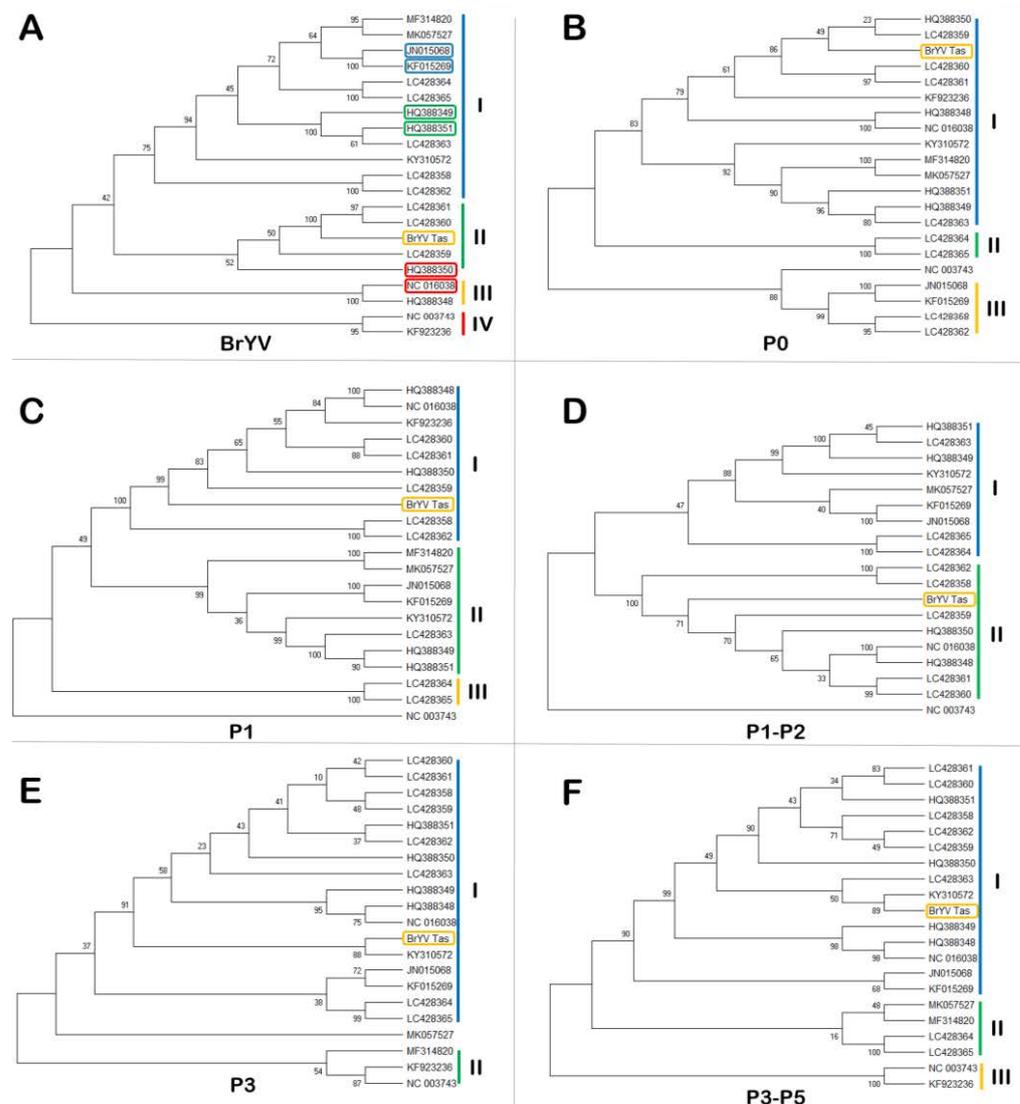


Figure 2. Cont.

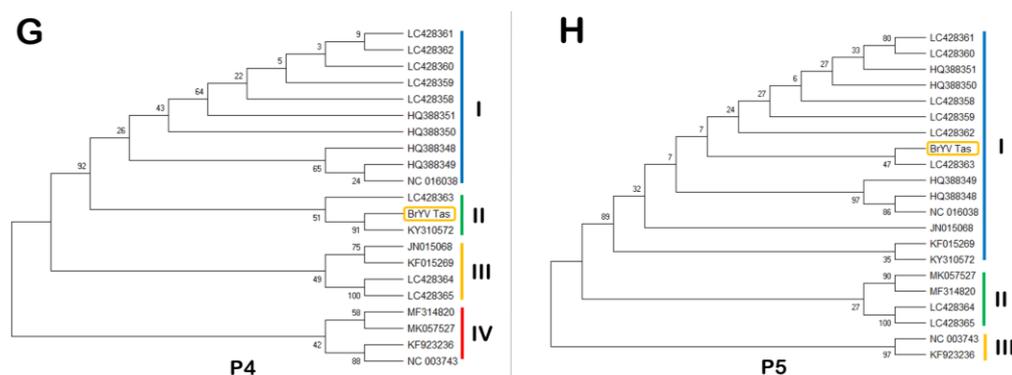


Figure 2. Phylogenetic trees (maximum likelihood algorithm) based on nucleotide sequence alignments of *Brassica Yellow Virus* (BrYV) constructed in MEGA-X. Bootstrap values shown at the nodes indicate the percentage of 1000 replications supporting the branching patterns shown. Phylogenetic analysis includes full/near-full genomes (A) and each gene responsible to encode their respective proteins (B) P0, (C) P1, (D) P1-P2, (E) P3, (F) P3-P5, (G) P4, and (H) P5. The accessions with red, green, and blue rectangles represent BrYV genotypes A, B, and C, respectively. Yellow rectangles indicate the BrYV-Tas isolate from the present study. Additional information regarding all these isolates is given in Table 2.

Table 2. BrYV and TuYV reference isolates retrieved from GenBank.

Isolate	Accession No.	Host	Country	References
BrYV-Tas	OM469309	<i>Raphanus raphanistrum</i>	Australia, Hobart	This study
BrYV-ABJ	HQ388348	<i>Brassica napus</i> var. <i>napobrassica</i>	China: Beijing	[2]
BrYV-BBJ	HQ388349	<i>B. napus</i> var. <i>napobrassica</i>	China: Beijing	[2]
BrYV-AJS	HQ388350	<i>B. campestris</i> L.	China: Jiangsu	[2]
BrYV-BJS	HQ388351	<i>B. campestris</i> L.	China: Jiangsu	[2]
BrYV-CR	JN015068	<i>R. raphanistrum</i>	China: Haidian	[6]
BrYV-CC	KF015269	<i>B. rapa pekinensis</i>	China: Beijing	[6]
BrYV-CS	KF923236	<i>B. rapa</i>	South Korea: Cheongsong	[3]
BrYV-China	KY310572	<i>B. napus</i>	China	Direct submission
BrYV-CC1	LC428358	<i>B. rapa</i> subsp. <i>pekinensis</i>	Japan: Hokkaido	[4]
BrYV-WN1	LC428359	<i>Sinapis alba</i>	Japan: Hokkaido	[4]
BrYV-TO3	LC428360	<i>B. rapa</i> subsp. <i>Rapa</i>	Japan: Okayama	[4]
BrYV-NAP	LC428361	<i>B. napus</i>	Japan: Okayama	[4]
BrYV-CD9	LC428362	<i>B. oleracea</i> var. <i>capitata</i>	Japan: Hokkaido	[4]
BrYV-R3b	LC428363	<i>R. sativus</i>	Japan: Okayama	[4]
BrYV-RT8	LC428364	<i>R. sativus</i>	Japan: Hokkaido	[4]
BrYV-R40	LC428365	<i>R. sativus</i>	Japan: Okayama	[4]
BrYV-Anhui	MF314820	<i>Nicotiana tabacum</i>	China	Direct submission
BrYV-NtabQJ	MK057527	<i>N. tabacum</i>	China: Yunnan	[17]
BrYV-ABJ	NC_016038	<i>B. napus</i> var. <i>napobrassica</i>	China: Beijing	[2]
TuYV-FL1	NC_003743	<i>Lactuca sativa</i>	France: Avignon	[20]

2.4. Recombination Analysis

When the full genomic sequences of different BrYVs were aligned, a significant recombination event for BrYV-Tas (recombinant), with LC428359 (major parent) and KY310572 (minor parent), was strongly supported using all seven methods in the RDP4 package, including RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq (Figure 3A). The p -values for these seven methods were as follows: RDP (3.103×10^{-11}), GENECONV (2.109×10^{-12}), BootScan (5.923×10^{-12}), MaxChi (1.002×10^{-15}), Chimaera (3.518×10^{-16}), SiScan (2.511×10^{-27}), and 3Seq (3.068×10^{-19}). We observed that the recombinant region covered almost 41% of the BrYV genome, and the predicted beginning and ending breakpoints are located at nt 1922 in the region of P1-P2 and nt 5234 in the P5 region, respectively (Figure 3B).



Figure 3. (A) Graphical representation of the recombinant region (pink). Pairwise identities of each pair of sequences (y-axis) and their position (x-axis) are indicated. (B) Schematic representation of BrYV-Tas isolate genome with the pattern of detected recombination breakpoints distributed in genome (highlighted).

2.5. Nucleotide Diversity and Haplotype Variability Indices

We analysed all datasets comprising BrYV full genomes and individual genes to compare the standing molecular diversity among viral populations arising from different geographical locations. The diversity determined for the whole genome analysis of BrYV was $S = 1477$, $\text{Eta} = 1684$, $\pi = 0.06883$, $H_d = 0.9950$, $\theta w = 0.07683$, Tajima's $D = -0.89807$ (Figure 4A and Table 3).

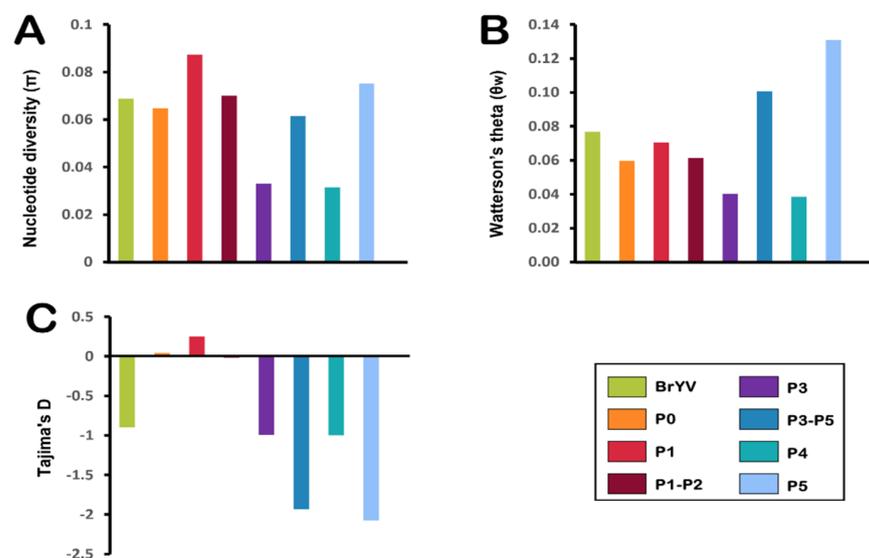


Figure 4. Estimation of genetic diversity was performed for the BrYV genome and individual genes. The calculated population genetic parameters include (A) nucleotide diversity (π), (B) Watterson's theta (θ_w), and (C) Tajima's D .

Table 3. Molecular diversity among BrYV populations and individual genes.

Dataset	Number of Sequences	Total Number of Sites	S	Eta	H	Hd	π	θw	Tajima's D
BrYV	20	5827	1477	1684	19	0.995	0.06883	0.07683	−0.89807
P0	20	751	159	171	19	0.995	0.06490	0.05976	0.04076
P1	20	1835	451	528	19	0.995	0.08736	0.07039	0.25116
P1-P2	20	3119	652	747	17	0.993	0.07001	0.06137	−0.01833
P3	20	609	87	94	18	0.989	0.03297	0.04027	−0.99362
P3-P5	20	2137	688	780	19	0.995	0.06144	0.10058	−1.93078
P4	20	528	72	78	14	0.932	0.03147	0.03844	−0.99753
P5	20	1342	506	578	19	0.995	0.07522	0.13073	−2.07541

S, number of polymorphic (segregating) sites; Eta, total number of mutations; H, number of haplotypes; Hd, haplotype diversity; π , nucleotide diversity; θw , Watterson's theta.

Genetic diversity analysis for BrYV individual genes also revealed that all genes were variable with high numbers of polymorphisms and polymorphic sites, very high haplotype diversity, but low nucleotide diversity (Figure 4A and Table 3). Aside from P3 (S = 87, Eta = 94, π = 0.03297, h = 0.989) and P4 (S = 72, Eta = 78, π = 0.03147, H_d = 0.932), each gene indicated high diversity and had a high proportion of haplotypes. P5 with Eta = 578, π = 0.07522, and H_d = 0.995 appeared to be the most genetically diverse gene that is three times more diverse than that of the conserved P3/P4 region of the BrYV genome (Table 3). Following P5, P1 (S = 451, Eta = 528, π = 0.08736, H_d = 0.995) appeared to be the gene with next greatest diversity (possibly linked to the genetically diverse region within P1) surpassing P0 (S = 159, Eta = 171, π = 0.06490, H_d = 0.995), previously regarded as the next most variable gene. Despite the fact that the polymorphic sites varied between 14 and 38% for each gene sequence, haplotype diversity was high, ranging from 93 to 100%, indicating that the number of unique sequences with different combinations of these polymorphic sites was high, resulting in greater diversity within the BrYV population (Table 3). Likewise, we also compared the genetic diversity between two genes encoding fusion proteins i.e., P1-P2 and P3-P5. Despite the existence of discrepancies between gene sizes, we were able to calculate average pairwise nucleotide diversities (π) for the aforementioned datasets (Figure 4A and Table 3).

The average number of segregating sites (θw) was greater for P5 (θw = 0.1307), followed by the P3-P5 (θw = 0.1006), the P1 (θw = 0.0704), the P1-P2 (θw = 0.0614), the P0 (θw = 0.0598), the P3 (θw = 0.0403), and the P4 (θw = 0.0384), respectively (Figure 4B and Table 3). The segregation rate exhibited by the overall BrYV population was θw = 0.0768. We also determined Tajima's D values in the BrYV populations and individual genes which appeared to be negative, ranging from −2.075 to −0.251, indicating the presence of excessive polymorphic sites among these isolates (Figure 4C and Table 3). Although the P0 and P1 displayed slightly positive Tajima's D values (0.0408 and 0.2512, respectively), this observation remained statistically non-significant.

2.6. Detection of Positive and Negative Selection Sites

We compared the non-synonymous to synonymous substitutions (d_N/d_S) for each gene to better understand the role of selection pressure on the genomic variation observed between the analysed datasets. DataMonkey-based analysis indicated that the genome of the BrYV population is being driven by purifying (negative) selection as the average d_N/d_S ratio of the BrYV genome remained <1 i.e., d_N/d_S = 0.902 (Figure 5 and Table 4).

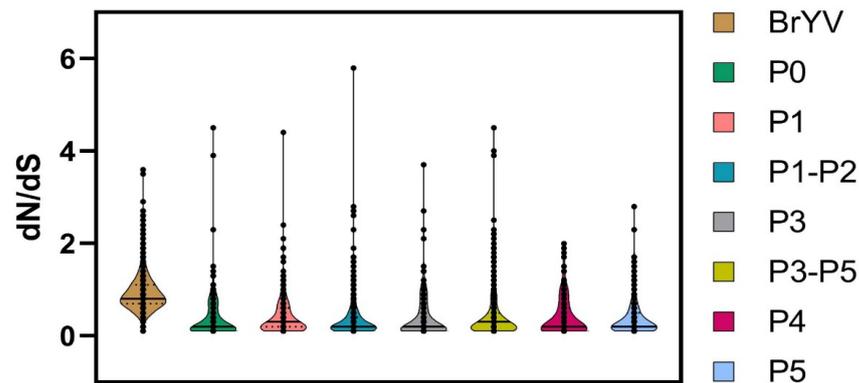


Figure 5. Estimation of selection pressure was performed by calculation of non-synonymous to synonymous substitution ratios (dN/dS). Violin plots correspond to dN/dS ratio for BrYV genome and its individual encoded genes. Solid horizontal lines represent median values, while dotted lines denote the lower and upper quartiles in each dataset.

Table 4. Estimation of average dN/dS ratios and positive and negative selection for coding sequences of BrYV.

ORF	Avg. dN/dS Ratio	Positive Selection				Negative Selection			
		Total Sites	Avg.	Min.	Max.	Total Sites	Avg.	Min.	Max.
BrYV	0.902	501	1.341	1.051	3.613	1167	0.699	0.051	0.949
P0	0.376	9	2.040	1.050	4.54	201	0.286	0.054	0.935
P1	0.414	22	1.548	1.063	4.401	282	0.319	0.050	0.942
P1-P2	0.369	42	1.534	1.051	5.829	679	0.289	0.050	0.937
P3	0.409	11	1.777	1.084	3.660	122	0.265	0.050	0.918
P3-P5	0.409	42	1.775	1.060	4.462	576	0.294	0.050	0.942
P4	0.402	11	1.487	1.098	2.026	127	0.280	0.052	0.932
P5	0.371	25	1.459	1.058	2.778	353	0.288	0.050	0.949

Overall, 1167/1848 sites were detected with negative selection, with values ranging between 0.051 and 0.949. Meanwhile, the proportion of positively selected sites remained lower (501/1848) than the negatively selected sites. Likewise, all genes displayed a variable frequency of negative selection with the lowest average dN/dS value for P5 (0.371) and the highest average value for P1 (0.414) (Figure 5 and Table 4). A total of 353/425 sites were found under negative selection in the P5 gene sequence with values ranging from 0.050 to 0.949. Positive selection was detected at a few sites (25/425) with values ranging between 1.058 and 2.778; however, the proportion remained much lower than for negatively selected sites (Figure 6 and Table 4).

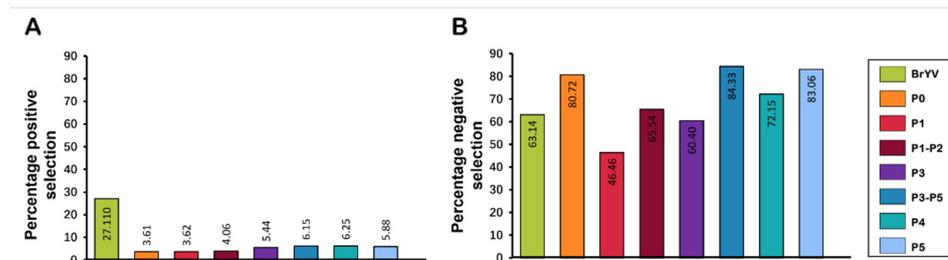


Figure 6. Percentage of sites evolving under negative/positive selection pressure. The datasets include (A) percentage of sites evolving under positive selection pressure in BrYV genome and its encoded genes and (B) percentage of negatively selected sites in BrYV genome and its encoded genes.

P1, with 282/607 negatively selected sites, displayed a similar pattern, having the average negative selection value of 0.319, with the lowest value of 0.050, and the highest of 0.942. Meanwhile, 22 sites were found to be positively selected in P1. With an average positive selection value of 1.548, the minimum value remained 1.063 and the maximum value was 4.401 (Table 4). P0 exhibited only nine positively selected sites with a value of 2.040, while this gene showed 201/249 negatively selected sites with 0.054 and 0.935 being the minimum and maximum values, respectively. P3 and P4 exhibited 11 positively selected sites, each with the ratio of the negatively selected sites being 122/202 and 127/176 for the P3 and P4 genes, respectively (Figure 6 and Table 4).

3. Discussion

The present study identified and genetically characterised an isolate of BrYV found infecting a *R. raphanistrum* host plant in Tasmania. Subsequent NGS data analysis revealed that this isolate of BrYV displayed a high percentage (96.68%) of nucleotide similarity with a BrYV isolate reported from Japan (LC428360). The NGS contig covered almost 70% of the genome of the BrYV reference sequence (NC_016038). A reconstructed genome, generated using NGS and RT-PCR, was 5516 bp long and had the typical organisation of BrYV. However, it was not possible to obtain the ~138 bp of the 3' end of the genome, which includes the latter half of 3' UTR. BrYV isolates from China, Korea, and Japan [2,3,6,17,21] have shown genetic variation in their ORFs and have been subdivided further into three genotypes (BrYV-A, B, and C) according to sequence comparisons and phylogenetic analysis [2,6]. This suggests that BrYV genomes, like those of other poleroviruses, are more diverse within different regions of the genome than previously thought, and that more diversity can be discovered after analysing whole-genome sequences. Our findings corroborate these previous studies [19], and multiplex RT-PCR analysis revealed that the BrYV-Tas isolate belongs to the BrYV genotype-A. Further analyses involving full genome and individual genes sequences of BrYV-Tas and other BrYV and TuYV isolates revealed that the presence of variation and clustering patterns were variable among different genes.

It has recently been proposed that TuYV and BrYV should be considered members of the same species if they share at least 83% nucleotide sequence identity [5]. The authors suggested a species demarcation criteria for new poleroviruses should include less than 83% nt sequence identity across all entire coding regions of the genome, as well as any significant differences in the host range, vector, and serology. The RdRp and CP regions were proposed as important for species demarcation with an 89% aa sequence identity in both regions for the members of the same polerovirus species [22], and CP is the most conserved region within poleroviruses [23,24]. However, currently, the working criterion for the demarcation of polerovirus species remains unchanged in the most recent report of ICTV on virus taxonomy [10], which is based on differences in host range, failure to cross-protect, differences in serological reactions, and differences in the amino acid sequence of any gene product greater than 10% [10,25]. Our findings are consistent with the recent species demarcation standards by ICTV [10] as well as with the previous studies related to characterisation of BrYV [2,3,6], as all genes of BrYV-Tas isolates shared relatively low (less than 90%) amino acid similarity with TuYV (NC_003743) with the exception of P3 and P4 genes which shared over 90% amino acid similarity with both BrYV and TuYV (Tables S12 and S14). Notably, for BrYV-Tas the cistrons at the 3' end shared the least sequence identity (60.9% for P3-P5, and 43.2% for P5) with TuYV. Such difference in the P5 sequence between the BrYV and the TuYV isolates could be linked to variability in aphid transmission specificity or efficiency [26]. Based on the ICTV report, nucleotide sequence analysis, and phylogeny results our findings provide compelling evidence that BrYV-Tas is distinct from TuYV.

We have shown evidence of recombination in the genome of BrYV-Tas, with the analysis suggesting that the recombinant consisted of genomic parts from distinct BrYV isolates. Recombination events in RNA viruses have been well known [27–29], and

both intraspecific, homologous and interspecific, non-homologous recombination are thought to be common and important in the evolution of poleroviruses. A few examples include BrYV [3], FBPV-1 [30], *Cucurbit aphid-borne yellows virus* (CABYV) [27,31], *Sugarcane yellow leaf virus* (ScYLV) [12,32], and *Cotton leafroll dwarf virus* (CLRDV) [33]. Our findings also suggest that recombination could play a role in the evolution of BrYV populations, as some of these genes, which are associated with compatible interactions, might be evolving under selection pressure from their hosts and/or vectors and tend to accumulate changes faster than other parts of the genome. Notably, mutations or recombination in genes can affect the biological functions of the viral proteins. However, how recombination could affect the biological functions associated with these genes remains a question worth studying in the future.

Our findings revealed that BrYV populations are evolving mainly under purifying selection pressure (Figures 5 and 6 and Table 4). To gain an in-depth understanding of this selection factor at the gene level, we opted to estimate d_N/d_S for P0, P1, P1-P2, P3, P3-P5, P4, and P5 genes of BrYV. While our results showed that the majority of the codons remained under negative selection for each gene with an average d_N/d_S ratio of <1 (Figure 5 and Table 4), the overall contribution of negatively selected sites remained more than the positively selected sites (Figure 6). Since all the sites had both positive and negative selection pressures, this provided strong evidence of heterogeneous selection pressures among codon sites. This is in accordance with a previous study that concluded that the genes of the TuYV population in the UK were evolving under negative selection pressure [34].

Future studies focusing on the patterns of BrYV geographical distribution, evolutionary dynamics, and functional analysis of the BrYV genes arising from genetic variations will help us to understand the host (plant/insect)-BrYV relationships and viral pathogenicity in the context of developing strategies for sustainable virus disease management.

4. Materials and Methods

4.1. Sample Collection and Nucleic Acid Extraction

In November 2019, ten samples of wild radish (*R. raphanistrum*) were randomly collected at a commercial vegetable farm in Deloraine as part of an extensive survey across Tasmania to determine the population diversity of *Polerovirus* species in pea crops and the weed species present in the vicinity of the pea crops. All leaf samples were grouped into lots of five and preserved (at $-80\text{ }^{\circ}\text{C}$) for subsequent testing. Extraction of total RNA was carried out using PowerLyzer 24 homogeniser (Qiagen, Chadstone, VIC, Australia) and RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Samples of 100 mg/lot were subject to RNA extraction to obtain sufficient yields of RNA for further processing and subsequent detection of viruses with low titer.

4.2. Next-Generation Sequencing (NGS) and Sequence Reads Analysis

Total RNA of the grouped sample 148 W was then sent to the Department of Jobs, Precincts and Regions, AgriBio, Bundoora, Victoria, for NGS. Libraries for RNA-Seq were prepared with a TruSeq Stranded Total RNA Sample Preparation Kit with Ribozero following the manufacturer's instructions. To determine the size distribution and concentration of the RNA-Seq libraries, 2200 TapeStation[®] system (Agilent Technologies, Santa Clara, CA, USA) and Qubit[®] Fluorometer 2.0 (Invitrogen) were used, respectively. Afterwards, the Illumina NovaSeq was employed to sequence these libraries with a paired read length of 2×151 bp.

Trimming of the primer and adaptor sequences from raw reads was carried out using Geneious Prime 2021.2 (Biomatters Ltd., Auckland, New Zealand). For this purpose, the BBDDuK plugin (provided in the BBTools package) was employed [35]. After that, the Velvet plugin (in Geneious Prime software) with default settings was used to perform de novo assembly. Then, a length-based sorting of the obtained contigs was carried out and, subsequently, blast (BLASTn and BLASTx) analysed against the GenBank database. Simultaneously, the

mapping of contigs and reads to a collection of polerovirus reference genomes (available in GenBank) was carried out and results were also analysed in Geneious Prime.

4.3. Validation of NGS Results and Amplification of BrYV Genome

To synthesize the cDNA, an iScript Reverse Transcription Supermix Kit (BioRad BioRad, Hercules, CA, USA) was used according to the provided instructions. The reaction mixture of 20 μ L contained total RNA (2.5 μ L), 5 \times RT Supermix (4 μ L), diethyl pyrocarbonate (DEPC), and treated water to make up the volume. The reaction was then incubated for priming at 22 $^{\circ}$ C (5 min), reverse transcription at 46 $^{\circ}$ C (20 min), and finally enzyme inactivation at 95 $^{\circ}$ C (1 min).

Several overlapping primer sets [17,18] were used to amplify the missing part of the virus genome, with reference to previously reported sequences of BrYV and TuYV (Table 1). This was done by RT-PCR-based amplification of the target sequence. The details, including primer sequences, amplicon size, target, and thermocycling conditions, are provided in Table 1. The assays were performed as follows: 1 \times HotStarTaq Master Mix (Qiagen), 10 μ M of each primer pair, 1 μ L cDNA template, and made up to a total volume of 20 μ L with sterile distilled water. Amplicon of appropriate sizes were then excised and purified using a QIAquick Gel Extraction Kit Protocol (Qiagen). Elution was carried out using a 30 μ L elution buffer (10 mM Tris HCl, pH 8.5) and sent for bidirectional (sanger) sequencing by the Central Science Laboratories (CSL), University of Tasmania, Australia. The near full-length genomic sequence of BrYV Tasmanian isolates (BrYV-Tas) was then deposited in the NCBI GenBank under accession number OM469309. In addition, multiplex RT-PCR with specific primer sets (Table 1) was also performed to determine the genotype of BrYV, as described previously [19].

4.4. Multiple Sequence Alignment and Phylogenetic Analysis

To prepare multiple sequence alignments, the near-full-length genome sequence of BrYV-Tas was aligned with the full genome sequences of all reported BrYV isolates and one TuYV isolate (NC_003743), available in the GenBank (Table 2), using the MUSCLE option in Geneious Prime 2021.2. Similarly, all genes including P0, P1, P1-P2, P3, P3-P5, P4, and P5 were also aligned with the corresponding genes of the reported isolates, as mentioned above. Maximum likelihood phylogenetic trees were constructed in MEGAX with 1000 bootstrap replicates [36].

4.5. Recombination Analysis

The aligned genome of the BrYV isolate from this study and other BrYV isolates available in GenBank were tested for potential recombination events. The recombination analysis was carried out using RDP [37], GENECONV [38], BOOTSCAN [39], MAXCHI [40], CHIMAERA [41], SISCAN [42], and 3SEQ [43] methods implemented in the recombination detection program v.4 [44]. Aligned sequences were subjected to the above-mentioned methods using default parameters. Potential recombination events detected were considered statistically significant if supported by at least four programs with p -values lower than a Bonferroni-corrected cutoff of 0.05.

4.6. Nucleotide Diversity and Haplotype Variability Indices

The average pairwise number of nucleotide differences per site (nucleotide diversity, π) was detected using DnaSP v.5 [45]. The statistically significant differences among the mean nucleotide diversity from the datasets were determined by calculating their 95% bootstrap confidence intervals. In addition, the nucleotide diversity was calculated using a 100-nucleotide sliding window, with a 10-nucleotide step size. The number of segregating sites (S), the number of haplotypes (H), and haplotype diversity (H_d) were also calculated for all datasets using DnaSP v.5 [45].

4.7. Detection of Positive and Negative Selection Sites

Using four different approaches, i.e., single-likelihood ancestor counting (SLAC), fixed-effects likelihood, random-effects likelihood, and partitioning for robust inference of selection [46], potential negatively and positively selected sites in the coding regions of P0, P1, P1-P2, P3, P3-P5, P4, and P5 were identified. All the analyses were performed in the Datamonkey web server (www.datamonkey.org accessed on 22 October 2021) [47]. Genetic Algorithm Recombination Detection (GARD) [48] was implemented to look for recombination breakpoints in the datasets in order to avoid ambiguous results. The dN/dS ratios were estimated using the SLAC method based on inferred GARD-corrected phylogenetic trees to compare the selection pressures acting on all genes of BrYV.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants11070884/s1>, Table S1: Percentage nucleotide sequence identities of whole genome of BrYV-Tas with other isolates reported worldwide; Table S2: Percentage nucleotide sequence identities of P0 gene of BrYV-Tas with other isolates reported worldwide; Table S3: Percentage nucleotide sequence identities of P1 gene of BrYV-Tas with other isolates reported worldwide; Table S4: Percentage nucleotide sequence identities of P1-P2 gene of BrYV-Tas with other isolates reported worldwide; Table S5: Percentage nucleotide sequence identities of P3 gene of BrYV-Tas with other isolates reported worldwide; Table S6: Percentage nucleotide sequence identities of P3-P5 gene of BrYV-Tas with other isolates reported worldwide; Table S7: Percentage nucleotide sequence identities of P4 gene of BrYV-Tas with other isolates reported worldwide; Table S8: Percentage nucleotide sequence identities of P5 gene of BrYV-Tas with other isolates reported worldwide; Table S9: Percentage amino acid sequence identities of P0 gene of BrYV-Tas with other isolates reported worldwide; Table S10: Percentage amino acid sequence identities of P1 gene of BrYV-Tas with other isolates reported worldwide; Table S11: Percentage amino acid sequence identities of P1-P2 gene of BrYV-Tas with other isolates reported worldwide; Table S12: Percentage amino acid sequence identities of P3 gene of BrYV-Tas with other isolates reported worldwide; Table S13: Percentage amino acid sequence identities of P3-P5 gene of BrYV-Tas with other isolates reported worldwide; Table S14: Percentage amino acid sequence identities of P4 gene of BrYV-Tas with other isolates reported worldwide; Table S15: Percentage amino acid sequence identities of P5 gene of BrYV-Tas with other isolates reported worldwide.

Author Contributions: Conceptualization, M.U. and C.R.W.; methodology, M.U., R.S.T., T.T. and C.R.W.; software, bioinformatics analysis, and validation, M.U. and T.F.; writing—original draft preparation, M.U.; writing—review and editing, R.S.T., T.F., T.T. and C.R.W.; supervision, R.S.T., T.T. and C.R.W.; project administration, C.R.W.; funding acquisition, C.R.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Hort Innovation, using the Vegetable Industry research and development levy and contributions from the Australian Government (project VG16086: Area wide management of vegetable diseases: virus and bacteria). Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is contained within the article or supplementary materials. The genome sequence of BrYV-Tas has been deposited in the NCBI database with Accession no: OM469309.

Acknowledgments: We thank the growers and Simplot Australia Pty Ltd. for access to the commercial field sites for sample collection and Cliff Kinoti, Agriculture Victoria Research, Department of Jobs, Precincts and Regions for providing the NGS services.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Garcia-Ruiz, H.; Holste, N.M.; LaTourrette, K. Poleroviruses (Luteoviridae). In *Encyclopedia of Virology*, 4th ed.; Bamford, D.H., Zuckerman, M., Eds.; Academic Press: Oxford, UK, 2021; pp. 594–602.
2. Xiang, H.-Y.; Dong, S.-W.; Shang, Q.-X.; Zhou, C.-J.; Li, D.-W.; Yu, J.-L.; Han, C.-G. Molecular characterization of two genotypes of a new polerovirus infecting brassicas in China. *Arch. Virol.* **2011**, *156*, 2251–2255. [[CrossRef](#)] [[PubMed](#)]
3. Lim, S.; Yoo, R.H.; Igori, D.; Zhao, F.; Kim, K.H.; Moon, J.S. Genome sequence of a recombinant brassica yellows virus infecting Chinese cabbage. *Arch. Virol.* **2015**, *160*, 597–600. [[CrossRef](#)] [[PubMed](#)]
4. Yoshida, N.; Tamada, T. Host range and molecular analysis of Beet leaf yellowing virus, Beet western yellows virus-JP and Brassica yellows virus in Japan. *Plant Pathol.* **2019**, *68*, 1045–1058. [[CrossRef](#)]
5. Filardo, F.; Nancarrow, N.; Kehoe, M.; McTaggart, A.R.; Congdon, B.; Kumari, S.; Aftab, M.; Trębicki, P.; Rodoni, B.; Thomas, J.; et al. Genetic diversity and recombination between turnip yellows virus strains in Australia. *Arch. Virol.* **2021**, *166*, 813–829. [[CrossRef](#)]
6. Zhang, X.Y.; Xiang, H.Y.; Zhou, C.J.; Li, D.W.; Yu, J.L.; Han, C.G. Complete genome sequence analysis identifies a new genotype of brassica yellows virus that infects cabbage and radish in China. *Arch. Virol.* **2014**, *159*, 2177–2180. [[CrossRef](#)]
7. Peter, K.A.; Gildow, F.; Palukaitis, P.; Gray, S.M. The C terminus of the polerovirus p5 readthrough domain limits virus infection to the phloem. *J. Virol.* **2009**, *83*, 5419–5429. [[CrossRef](#)]
8. Simon-Loriere, E.; Holmes, E.C. Why do RNA viruses recombine? *Nat. Rev. Microbiol.* **2011**, *9*, 617–626. [[CrossRef](#)]
9. Sztuba-Solińska, J.; Urbanowicz, A.; Figlerowicz, M.; Bujarski, J.J. RNA-RNA recombination in plant virus replication and evolution. *Annu. Rev. Phytopathol.* **2011**, *49*, 415–443. [[CrossRef](#)]
10. Walker, P.J.; Siddell, S.G.; Lefkowitz, E.J.; Mushegian, A.R.; Adriaenssens, E.M.; Alfenas-Zerbini, P.; Davison, A.J.; Dempsey, D.M.; Dutilh, B.E.; García, M.L. Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses. *Arch. Virol.* **2021**, *166*, 2633–2648. [[CrossRef](#)]
11. Beuve, M.; Stevens, M.; Liu, H.-Y.; Wintermantel, W.M.; Hauser, S.; Lemaire, O. Biological and Molecular Characterization of an American Sugar Beet-Infecting Beet western yellows virus Isolate. *Plant Dis.* **2008**, *92*, 51–60. [[CrossRef](#)]
12. Moonan, F.; Molina, J.; Mirkov, T.E. Sugarcane yellow leaf virus: An emerging virus that has evolved by recombination between luteoviral and poleroviral ancestors. *Virology* **2000**, *269*, 156–171. [[CrossRef](#)] [[PubMed](#)]
13. Stuart, G.W.; Moffett, P.K.; Bozarth, R.F. A comprehensive open reading frame phylogenetic analysis of isometric positive strand ssRNA plant viruses. *Arch. Virol.* **2006**, *151*, 1159–1177. [[CrossRef](#)] [[PubMed](#)]
14. Fox, A.; Mumford, R.A. Plant viruses and viroids in the United Kingdom: An analysis of first detections and novel discoveries from 1980 to 2014. *Virus Res.* **2017**, *241*, 10–18. [[CrossRef](#)] [[PubMed](#)]
15. Kehoe, M.A.; Coutts, B.A.; Buirchell, B.J.; Jones, R.A. Plant virology and next generation sequencing: Experiences with a Potyvirus. *PLoS ONE* **2014**, *9*, e104580. [[CrossRef](#)]
16. Roossinck, M.J.; Martin, D.P.; Roumagnac, P. Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology* **2015**, *105*, 716–727. [[CrossRef](#)]
17. Wang, Q.; Xu, F.Z.; An, L.L.; Xiang, H.Y.; Zhang, W.H.; Liu, G.S.; Liu, H.B. Molecular characterization of a new recombinant brassica yellows virus infecting tobacco in China. *Virus Genes* **2019**, *55*, 253–256. [[CrossRef](#)]
18. Schubert, J.; Rabenstein, F.; Graichen, K.; Richter, K. Comparison of the 5′ end nucleotide sequences of luteoviruses from oilseed rape and sugar beet. *Arch. Phytopathol. Plant Prot.* **1998**, *31*, 519–530. [[CrossRef](#)]
19. Zhang, X.; Peng, Y.; Wang, Y.; Zhang, Z.; Li, D.; Yu, J.; Han, C. Simultaneous detection and differentiation of three genotypes of Brassica yellows virus by multiplex reverse transcription-polymerase chain reaction. *Virol. J.* **2016**, *13*, 189. [[CrossRef](#)]
20. Veidt, I.; Lot, H.; Leiser, M.; Scheidecker, D.; Guillely, H.; Richards, K.; Jonard, G. Nucleotide sequence of beet western yellows virus RNA. *Nucleic Acids Res.* **1988**, *16*, 9917–9932. [[CrossRef](#)]
21. Kamitani, M.; Nagano, A.J.; Honjo, M.N.; Kudoh, H. RNA-Seq reveals virus–virus and virus–plant interactions in nature. *FEMS Microbiol. Ecol.* **2016**, *92*, fiw176. [[CrossRef](#)]
22. Kuhn, J.H.; Wolf, Y.I.; Krupovic, M.; Zhang, Y.-Z.; Maes, P.; Dolja, V.V.; Koonin, E.V. Classify viruses—The gain is worth the pain. *Nature* **2019**, *566*, 318–320. [[CrossRef](#)] [[PubMed](#)]
23. Fiallo-Olivé, E.; Navas-Hermosilla, E.; Ferro, C.G.; Zerbini, F.M.; Navas-Castillo, J. Evidence for a complex of emergent poleroviruses affecting pepper worldwide. *Arch. Virol.* **2018**, *163*, 1171–1178. [[CrossRef](#)] [[PubMed](#)]
24. Mayo, M.; d’Arcy, C. Family Luteoviridae: A reclassification of luteoviruses. In *The Luteoviridae*; Smith, H.G., Barke, R.H., Eds.; CABI Publishing: Wallingford, UK, 1999; pp. 15–22.
25. King, A.M.; Lefkowitz, E.; Adams, M.J.; Carstens, E.B. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*; Elsevier: Amsterdam, The Netherlands, 2011; Volume 9.
26. Stevens, M.; Freeman, B.; LIU, H.Y.; Herrbach, E.; Lemaire, O. Beet poleroviruses: Close friends or distant relatives? *Mol. Plant Pathol.* **2005**, *6*, 1–9. [[CrossRef](#)] [[PubMed](#)]
27. Costa, T.M.; Blawid, R.; Aranda, M.A.; Freitas, D.M.S.; Andrade, G.P.; Inoue-Nagata, A.K.; Nagata, T. Cucurbit aphid-borne yellows virus from melon plants in Brazil is an interspecific recombinant. *Arch. Virol.* **2019**, *164*, 249–254. [[CrossRef](#)]
28. Garcia-Ruiz, H.; Diaz, A.; Ahlquist, P. Intermolecular RNA recombination occurs at different frequencies in alternate forms of brome mosaic virus RNA replication compartments. *Viruses* **2018**, *10*, 131. [[CrossRef](#)] [[PubMed](#)]

29. Nagy, P.D. Recombination in plant RNA viruses. In *Plant Virus Evolution*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 133–156.
30. Filardo, F.F.; Thomas, J.E.; Webb, M.; Sharman, M. Faba bean polerovirus 1 (FBPV-1); a new polerovirus infecting legume crops in Australia. *Arch. Virol.* **2019**, *164*, 1915–1921. [[CrossRef](#)] [[PubMed](#)]
31. Knierim, D.; Deng, T.C.; Tsai, W.S.; Green, S.K.; Kenyon, L. Molecular identification of three distinct Polerovirus species and a recombinant Cucurbit aphid-borne yellows virus strain infecting cucurbit crops in Taiwan. *Plant Pathol.* **2010**, *59*, 991–1002. [[CrossRef](#)]
32. Moonan, F.; Mirkov, T.E. Analyses of genotypic diversity among North, South, and Central American isolates of sugarcane yellow leaf virus: Evidence for Colombian origins and for intraspecific spatial phylogenetic variation. *J. Virol.* **2002**, *76*, 1339–1348. [[CrossRef](#)]
33. Silva, T.F.; Corrêa, R.L.; Castilho, Y.; Silvie, P.; Bélot, J.L.; Vaslin, M.F. Widespread distribution and a new recombinant species of Brazilian virus associated with cotton blue disease. *Virol. J.* **2008**, *5*, 123. [[CrossRef](#)]
34. Newbert, M.J. The Genetic Diversity of Turnip Yellows Virus in Oilseed Rape (*Brassica Napus*) in Europe, Pathogenic Determinants of Resistance and Host Range. Ph.D. Thesis, University of Warwick, Coventry, UK, 2016.
35. Bushnell, B. BBTools: A Suite of Bioinformatic Tools Used for DNA and RNA Sequence Data Analysis. 2016.
36. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
37. Martin, D.; Rybicki, E. RDP: Detection of recombination amongst aligned sequences. *Bioinformatics* **2000**, *16*, 562–563. [[CrossRef](#)] [[PubMed](#)]
38. Sawyer, S. GENECONV: A Computer Package for the Statistical Detection of Gene Conversion. 1999. Available online: <http://www.math.wustl.edu/~sawyer> (accessed on 29 October 2021).
39. Salminen, M.O.; Carr, J.K.; Burke, D.S.; McCutchan, F.E. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res. Hum. Retrovir.* **1995**, *11*, 1423–1425. [[CrossRef](#)] [[PubMed](#)]
40. Smith, J.M. Analyzing the mosaic structure of genes. *J. Mol. Evol.* **1992**, *34*, 126–129. [[CrossRef](#)] [[PubMed](#)]
41. Posada, D.; Crandall, K.A. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13757–13762. [[CrossRef](#)] [[PubMed](#)]
42. Gibbs, M.J.; Armstrong, J.S.; Gibbs, A.J. Sister-Scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* **2000**, *16*, 573–582. [[CrossRef](#)] [[PubMed](#)]
43. Boni, M.F.; Posada, D.; Feldman, M.W. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **2007**, *176*, 1035–1047. [[CrossRef](#)]
44. Martin, D.P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, vev003. [[CrossRef](#)]
45. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [[CrossRef](#)]
46. Scheffler, K.; Martin, D.P.; Seoighe, C. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* **2006**, *22*, 2493–2499. [[CrossRef](#)]
47. Pond, S.L.K.; Frost, S.D.W. Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **2005**, *21*, 2531–2533. [[CrossRef](#)]
48. Kosakovsky Pond, S.L.; Posada, D.; Gravenor, M.B.; Woelk, C.H.; Frost, S.D.W. GARD: A genetic algorithm for recombination detection. *Bioinformatics* **2006**, *22*, 3096–3098. [[CrossRef](#)] [[PubMed](#)]