Article

Genetic Diversity of *Rice stripe necrosis virus* and New Insights into Evolution of the Genus *Benyvirus*

Issiaka Bagayoko [1,†], Marcos Giovanni Celli [2,3,†], Gustavo Romay [1], Nils Poulicard [4], Agnès Pinel-Galzi [4], Charlotte Julian [4,5], Denis Filloux [4,5], Philippe Roumagnac [4,5], Drissa Sérémé [6], Claude Bragard [1] and Eugénie Hébrard [4,*]

1   Earth and Life Institute, Applied Microbiology-Phytopathology, Université Catholique de Louvain (UCLouvain), Croix du Sud 2 Bte L07.05.03, 1348 Louvain-la-Neuve, Belgium; issiaka.bagayoko@uclouvain.be (I.B.); gromay@gmail.com (G.R.); claude.bragard@uclouvain.be (C.B.)
2   Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires 1425, Argentina; marcoscelli@hotmail.com
3   Instituto de Patología Vegetal (IPAVE, CIAP, INTA), Camino 60 cuadras Km 5, Cordoba 5119, Argentina
4   PHIM, Plant Health Institute, Université de Montpellier, IRD, INRAE, CIRAD, SupAgro, 911 Avenue Agropolis, 34394 Montpellier, France; nils.poulicard@ird.fr (N.P.); agnes.pinel@ird.fr (A.P.-G.); charlotte.julian@cirad.fr (C.J.); denis.filloux@cirad.fr (D.F.); philippe.roumagnac@cirad.fr (P.R.)
5   CIRAD, UMR PHIM, Campus International de Montferrier-Baillarguet, 34398 Montpellier, France
6   Laboratoire de Laboratoire de Virologie et de Biotechnologies Végétales, INERA—Institut de l'Environnement et de Recherches Agricoles, LMI Patho-Bios, Ouagadougou 01 BP 476, Burkina Faso; drissa.sereme@coraf.org
*   Correspondence: eugenie.hebrard@ird.fr
†   These authors contributed equally to this work.

Abstract: The rice stripe necrosis virus (RSNV) has been reported to infect rice in several countries in Africa and South America, but limited genomic data are currently publicly available. Here, eleven RSNV genomes were entirely sequenced, including the first corpus of RSNV genomes of African isolates. The genetic variability was differently distributed along the two genomic segments. The segment RNA1, within which clusters of polymorphisms were identified, showed a higher nucleotidic variability than did the beet necrotic yellow vein virus (BNYVV) RNA1 segment. The diversity patterns of both viruses were similar in the RNA2 segment, except for an in-frame insertion of 243 nucleotides located in the RSNV tgbp1 gene. Recombination events were detected into RNA1 and RNA2 segments, in particular in the two most divergent RSNV isolates from Colombia and Sierra Leone. In contrast to BNYVV, the RSNV molecular diversity had a geographical structure with two main RSNV lineages distributed in America and in Africa. Our data on the genetic diversity of RSNV revealed unexpected differences with BNYVV suggesting a complex evolutionary history of the genus *Benyvirus*.

Keywords: RSNV; benyvirus; diversity; recombination

1. Introduction

Rice is one of the major staple foods for global consumption and plays a crucial role in food security in developing countries [1]. More than 15 viruses have been reported to exist around the world, causing high yield losses. While most of these viruses are distributed in Asia, only two have recurrently been reported in Africa and America, i.e., the rice yellow mottle virus (RYMV) in Africa and the rice hoja blanca virus (RHBV) in America. However, other rice viruses have spread to those two continents. For instance, the rice stripe necrosis virus (RSNV) has recently emerged in America and is re-emerging in Africa. Hence, while RSNV was initially identified in the Ivory Coast in 1983 [2] and almost undetected since then, characteristic symptoms of crinkling yellow and deformation of rice leaves were recently observed in West-African countries, including Burkina Faso [3], Benin [4], Mali [5] and Sierra Leone [6]. The incidence of disease ranges locally from 37% to 80% in these

African countries. In South and Central America, the virus has been episodically reported in several countries, including Colombia, Brazil, Ecuador, Panama and Argentina [7–11], causing severe epidemics leading to up to 40% yield losses [10].

Rice stripe necrosis virus belongs to the genus *Benyvirus*, family *Benyviridae* [12]. The type member of the genus, beet necrotic yellow vein virus (BNYVV), causes rhizomania, one of the most important diseases affecting sugar beet [13]. Distributed worldwide, BNYVV is responsible for yield losses of up to 50%. In addition to RSNV and BNYVV, two other benyviruses have been characterized so far, beet soil-borne mosaic virus (BSBMV), present in the USA in sugar beet, but without inducing symptoms on roots [14,15] and burdock mottle virus (BdMV) described in Japan in the case of burdock plants (*Arctium lappa*) [16].

Benyviruses have non-enveloped multipartite rod-shaped particles. They are transmitted by soil-borne plasmodiophorid protists of the genus *Polymyxa*. Benyviruses can persist in the soil for decades through the vector spores [17]. Benyviruses have positive single-stranded RNA genomes composed of two to five segments, depending on the viral species [12]. The housekeeping genes are encoded by the two largest segments of ca. 6600 and 4600 nucleotides, respectively. RNA1 contains a single large open reading frame (ORF) that encodes the replication polyprotein including methyltransferase (Mtr), helicase (Hel), protease (Pro) and the RNA dependent RNA polymerase (RdRp) domains [12]. RNA2 contains six ORFs (ORF2-ORF7). ORF2 encodes the major coat protein (CP). ORF3 is translated from a readthrough (RT) mechanism producing a fusion of the CP and RT domain embedded in the viral capsid and participates in the vector transmission [12]. ORF4-6 form the Triple Gene Block (TGB) involved in virus movement. TGBp1 binds to RNA and interacts with TGBp2 and TGBp3 [18]. ORF7 encodes the cysteine-rich protein (CRP) dedicated to the viral suppressor of RNA silencing (VSR) [19]. BNYVV and BSBMV clones have been constructed using RNA1 and RNA2 segments and have the ability to infect *Nicotiana benthamiana* and *Beta macrocarpa* [20,21]. However, in nature, BNYVV, BSBMV and BdMV have additional small genomic components (RNA3 and RNA4). RNA3 contains two ORFs that encode a p25 protein responsible for the modulation of symptoms in leaves and roots, and a 6.8 kDa protein expressed only from RNA3 truncated forms [22,23]. RNA4 encodes a p31 protein, described as a root specific VSR [20], and is involved in vector transmission [20,24]. RNA5 is only present within some BNYVV isolates, and encodes a p26 protein, which induces severe rhizomania symptoms and highly-reduced sugar yields [25]. It has been suggested that RNA5 originates from a duplication of RNA3 [26].

More than 110 isolates of BNYVV have been partly sequenced, and several genetic groups have been described [27–29]. Molecular analyses have been mainly focused on small segments 3, 4 and 5 and on the *cp* gene (RNA2). Up to now, only nine entire sequences of BNYVV RNA1 and RNA2 are publicly available. This lack of knowledge originated from the major role of the small segments in rhizomania symptoms, but also on the low diversity observed from the nine RNA1 and RNA2 sequences [30]. However, the housekeeping genes contained by both RNA1 and RNA2 segments are crucial, and even sufficient for RSNV, to fulfil the viral cycle of benyviruses. Until 2019, only a single whole RSNV genome (RNA1+RNA2) of a Colombian isolate was made publicly available in GenBank. Recently, four additional genomes from Brazil have been released, likely linked to the emergence of RSNV in South America, and to the advances in high-throughput sequencing (HTS) technologies. In addition, our team released two sequences from Mali. The aim of this study was to fill the knowledge gap on the genetic diversity of RSNV using a combination of classical Sanger and Illumina sequencing. As a result, we obtained nine new entire genomes. The sequences from Burkina Faso, Sierra Leone and Benin constitute with those from Mali, the first corpus of African genomic sequences. This large amount of data has allowed new genetic diversity analyses and comparisons at intra- and inter-specific levels, revealing unexpected features of *Benyvirus*.

2. Materials and Methods

2.1. Viral Isolates

Seven samples were collected between 2016 and 2018 in Argentina (SF2.1 and Cor18), Burkina Faso (BF01Bama, BF02Bama and BF03Bama) and Mali (ML01San and ML02San) (Table 1). The isolates from Argentina originated from irrigated rice plots in Corrientes Province and from pluvial rice fields in the Santa Fe Province. The isolates from Mali originated from lowland rice fields in the San region, and from pluvial rice plots at Bama in Burkina Faso. Four rice samples previously collected between 2014 and 2019 from Benin (Be2) [4], Mali (M1-1 and M2-1) [5] and Sierra Leone (SL254) [6] were also used in this study. Specifically, Malian isolates have been maintained under greenhouse conditions in separate tube cultures of rice var. Kogoni 91-1 infected with its vector *P. graminis*.

Table 1. List of RSNV isolates used in this study with origin, collection date, accession numbers and bibliographic reference.

Country	Date	Name	Accession Numbers	Reference
Argentina	2017	Cor4	Hel MG792545	[11]
Argentina	2017	STFe5	Hel MG792546	[11]
Argentina	2017	STFe10.1	Hel MG792544	[11]
Argentina	2017	SF2.1	ARN1 MW147224, ARN2 MW147225	This study
Argentina	2018	Cor18	ARN1 MW147222, ARN2 MW147223	This study
Benin	2013	Be1	Hel KP099623	[4]
Benin	2013	Be2	ARN1 MW187590, ARN2 MW187591	This study
Brazil	2019	UFT/2019	ARN1 MT027255, ARN1 MT027256	Direct submission
Brazil	2019	BR-LE01-19	ARN1 MT270127, ARN2 MT270128	Direct submission
Brazil	2019	BR-TT01-19	ARN1 MT270129, ARN2 MT270130	Direct submission
Brazil	2020	UFGEm/2020-1	ARN1 MT507288, ARN2 MT507289	Direct submission
Burkina Faso	2013	BFRS1	CP LK023710	[3]
Burkina Faso	2016	BF-VK1	Hel MF115599, CP MF115604	[5]
Burkina Faso	2016	BF-K1	Hel MF115600, CP MF115605	[5]
Burkina Faso	2016	BF01Bama	ARN1 MW117944, ARN2 MW117939	This study
Burkina Faso	2016	BF02Bama	ARN1 MW117945, ARN2 MW117940	This study
Burkina Faso	2016	BF03Bama	ARN1 MW117946, ARN2 MW117941	This study
Colombia	2007	col	ARN1 EU099844, ARN2 EU099845	[9]
Mali	2015	MALI-SK1	Hel MF115602, CP MF115607	[5]
Mali	2015	MALI-I1	Hel MF115603, CP MF115608	[5]
Mali	2016	MALI-B1	Hel MF115601, CP MF115606	[5]
Mali	2016	M1-1	ARN1 MK170452, ARN2 MK170454	Direct submission
Mali	2016	M2-1	ARN1 MK170453, ARN2 MK170455	Direct submission
Mali	2016	ML01San	ARN1 MW117948, ARN2 MW117943	This study
Mali	2016	ML02San	ARN1 MW117947, ARN2 MW117942	This study
Sierra Leona	2019	SL254	Hel MN750254, CP MN750255, ARN1 MW187592, ARN2 MW187593	[6] This study

2.2. Genomic Sequencing

Different strategies were used to obtain the genomic sequences of the RSNV isolates. Total RNA extractions were performed from leaves or roots using the RNeasyPlant Mini kit (Qiagen) or TRIzol reagent (Invitrogen), according to manufacturers' instructions. The sequence from Benin was amplified using primers at the 5' and 3' ends of the two RNA segments (Table S1) followed by library constructs according to the protocol recommended in [31]. Sequencing was performed by Illumina, and assembly was performed using Galaxy workflows [32,33], implemented on the South Green bioinformatic platform as described in [34]. The sequence Cor18 from Argentina was obtained using VANA-NGS (Virion-Associated Nucleic Acids—Next-Generation Sequencing, [35,36]) sequenced by Illumina, and assembled in Geneious software v.9.1.8 (Biomatters Limited, Auckland, NewZealand). The genomes of the other nine isolates were amplified by RT-PCR using different combinations of specific primers (Table S1), followed by sequencing by the Sanger method, and assembled using BioEdit software [37]. For the isolates from Burkina Faso and Mali, PCR products from the 5' and 3' ends of the genome were sequenced after cloning into pGEM-T Easy Vector (Promega, Madison, WI, USA), while internal amplicons were directly sequenced by the Sanger method.

2.3. Sequence Analyses

The eleven RNSV sequences obtained in this study, the reference genome from Colombia, and the four newly-released sequences from Brazil and available in the GenBank database [38] (Table 1) were aligned using MUSCLE in SEAVIEW v4.7 software [39]. Conserved functional domains of benyviruses were located along the sequences [40]. Nine BNYVV entire publicly available sequences were used for diversity comparison with RSNV (accession numbers RNA1 and RNA2, respectively: HM126464 and HM11790, MH106726 and MH106727, KX665536 and KX665537, KM434313 and KM434314, EU330453 and EU330452, EU330450 and EU330451, X05147 and X04197, D84410 and D84411, MT227164 and MT227165). The location of hydrophobic clusters and predicted disordered domains in TGBp1 were analyzed using HCA v1.0.2 (https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py accessed on 18 September 2020) [41] and PONDR® VL-XT (http://pondr.com accessed on 18 September 2020) [42,43], respectively. The presence of nucleolar localization signals (NoLS) and nuclear localization signals (NLS) was searched for in TGBp1 using a nucleolar localization sequence detector (NOD) [44] and a cNLS mapper, available online for prediction of NLS specific to the importin $\alpha\beta$ pathway [45]. Helicase conserved motifs I-VI were located in the TGBp1 of RSNV, BNYVV, BSBMV, BdMV, potato mop-top virus (PMTV), peanut clump virus (PCV) and barley stripe mosaic virus (BSMV) (accession numbers: EU099845, MK170454, MT227165, JF513083, AB818899, MK539869, L07269 KY615805) based on [46]. Diversity analyses were performed independently for RSNV and BNYVV using DNAsp software, version 6.12.03 with default parameters [47] (Table S2). The location of parsimony informative polymorphisms was observed along the genome. Putative recombination events were identified using seven methods (RDP, GeneConv, Chimaera, MaxChi, BootScan, SiScan, and 3Seq) implemented in the RDP4 software, version 4.974 [48]. Only recombination events detected by at least three methods were considered and sequence alignments with recombinant regions removed were generated for phylogenetic analyses. Maximum-likelihood (ML) phylogenetic trees were reconstructed using the PHYML-3.1 algorithm implemented in SEAVIEW v4.7 [39] under the best-fitted nucleotide substitution model (TN93+G, K2+G, K2+I for the complete genome, the helicase and the CP datasets, respectively), with gap treatment by pairwise deletion and with 100 bootstrap replications. Phylogenetic trees were drawn using FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/ accessed on 17 September 2020).

3. Results

To date, five RSNV genomes have been entirely sequenced: one from the reference isolate collected in 1999 in Colombia [7], and four from isolates collected in 2019–2020

in Brazil. In this study, eleven new RSNV isolates were fully sequenced including the sequences from Mali recently released (Table 1). Two isolates from Argentina were collected in the major rice producing areas in Corrientes and Santa Fe Provinces. Nine isolates were originated from Mali, Burkina Faso, Benin and Sierra Leone. This dataset constitutes the first corpus of African RNSV genomic sequences.

All RSNV full-length sequences showed the genomic organization of benyviruses (Figure 1). RNA1 segment encodes the replication polyprotein including Mtr, Hel, Pro and RdRp. RNA2 segment harbors six open reading frames encoding: CP, RTD, TGBp1, TGBp2, TGBp3 and CRP.
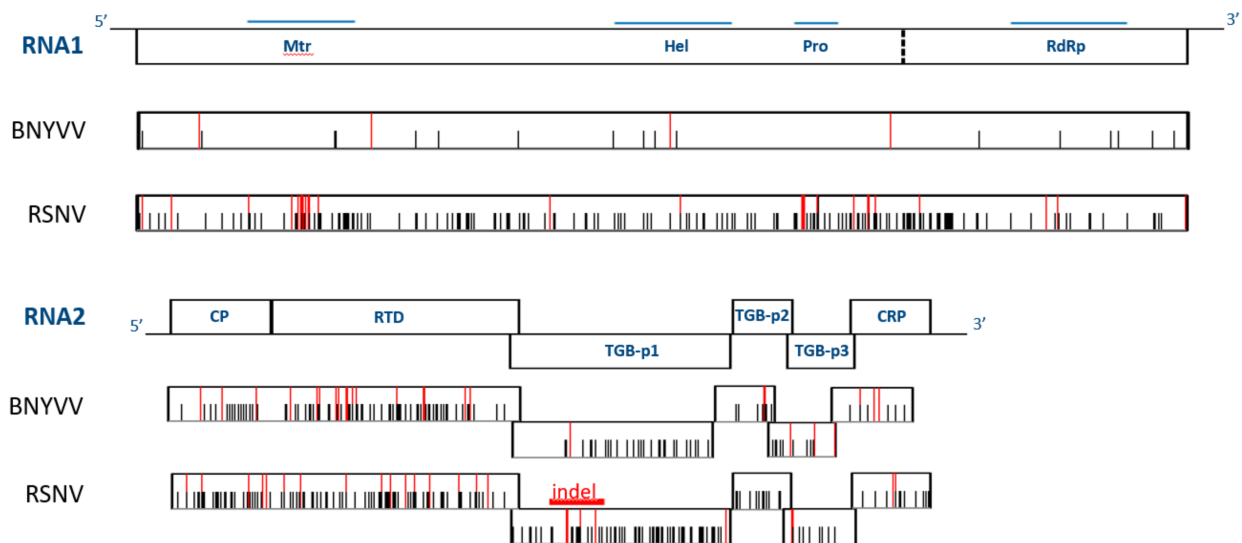
Figure 1. Polymorphisms along the RNA1 and RNA2 segments of RSNV and BNYVV. Black horizontal lines indicate untranslated regions, rectangular boxes indicate ORFs and blue horizontal lines indicate conserved functional domains. Black and red vertical bars indicate positions of parsimony informative polymorphisms (synonymous and non-synonymous polymorphisms, respectively), based on DNAsp analyses independently for each viral species [47]. Red horizontal line indicates the insertion in RNSV TGBp1. Mtr, methyltransferase, Hel, helicase, Pro, protease, RdRp, RNA-dependent RNA polymerase, CP, coat protein, RTD, read-through domain, TGBp, triple gene block protein, CRP, cystein-rich protein.

There were no insertion-deletion events in RNA1. Compared to the publicly available sequences, the RNA2 of all the new isolates showed a longer sequence due to an in-frame insertion of 243 nucleotides length at the position 252 of the tgbp1 gene (corresponding to position 2444 in the RNA2 segment). This insertion was also observed in the UFGEm/2020-1 sequence from Brazil. The 81 amino acids inserted into the N-terminal part of TGBp1 did not show any homology with sequences available in the GenBank database, or in known conserved domains (data not shown). The analysis of the hydrophobic amino acid clusters revealed that this domain was absent from the BNYVV TGBp1 (Figure S1). The insertion contains a low number of hydrophobic amino acid clusters, a high number of charged residues (19), and patches of glycine (7), serine (12) and proline (5). These features would result in a second disordered domain in the RSNV TGBp1, as revealed by the PONDR® VL-XT analysis (Figure S2). Moreover, in addition to a NLS found in all RSNV TGBp1, a NoLS has been detected exclusively in the additional sequence (Figure 2).
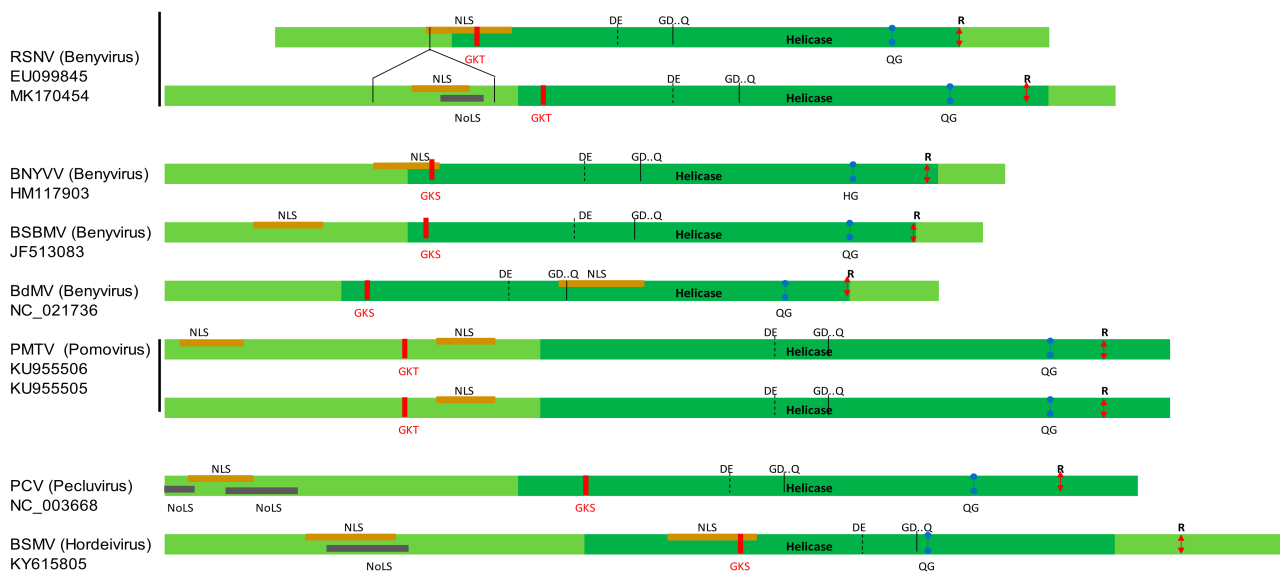
Figure 2. Organization of benyvirus and hordei-like TGBp1. Nucleolar localization signal (NoLS) and nuclear localization signals (NLS) were detected using Nucleolar localization sequence detector (NOD) [44] and cNLS mapper, for prediction of NLS specific to the importin αβ pathway [45]. Helicase domain and its conserved motifs I-VI were indicated in dark green based on [46]. The RSNV insertion domain is indicated by black lines. RSNV, rice stripe necrosis virus isolates col and M1-1, BNYVV, beet necrotic yellow vein virus, BSBMV, beet soil-borne mosaic virus, BdMV, burdock mottle virus, PMTV, potato mop-top virus, PCV, peanut clump virus, BSMV, barley stripe mosaic virus.

The RSNV sequences were highly conserved (>97.5% sequence identity). However, the overall mean distances of the 16 RSNV genomes were higher than expected in comparison to the nine BNYVV genomes (0.023 subst/nt $\pm$ 0.002 and 0.025 subst/nt $\pm$ 0.002 for RSNV RNA1 and RNA2 compared to 0.007 subst/nt $\pm$ 0.001 and 0.014 subst/nt $\pm$ 0.004 for BNYVV RNA1 and RNA2, respectively). It is worth noting that the value from RSNV RNA1 was three times higher than from BNYVV. The genetic variability was differentially distributed along the two viral segments (Figure 1). In RSNV RNA1, genetic regions of Mtr, of Pro and between Pro and RdRp showed the highest variability at synonymous and non-synonymous levels. Significantly positive Tajima's D values were observed in these domains (Figure S3). The genetic diversity of RSNV and BNYVV RNA2 showed less contrasted patterns. Synonymous and non-synonymous substitutions were homogeneously distributed in the cp-rt gene. The number of non-synonymous changes was low in the tgbp and crp genes.

Recombination detection was performed using the RDP4 software (Figure 3, Table 2). For the first time in the genus *Benyvirus*, statistically significant evidence for recombination was detected in both segments ($3.42 \times 10^{-37} < P < 4.67 \times 10^{-4}$). Four and two putative recombination events were identified in RNA1 and RNA2, respectively. In RNA1, putative recombinant events were located in the Mtr region (isolate M1-1), between Mtr and Pro regions (isolate SL254) and between the Pro and RdRp regions (isolates col & Be2). In RNA2, the putative recombinant events concerned one or the other part of the two-thirds of the segment (isolates Be2, BR-LE01-19 and col). The same event RNA2#2 was detected in isolates BR-LE01-19 and col. One of the two putative parents was still unknown in 4/6 recombination events.
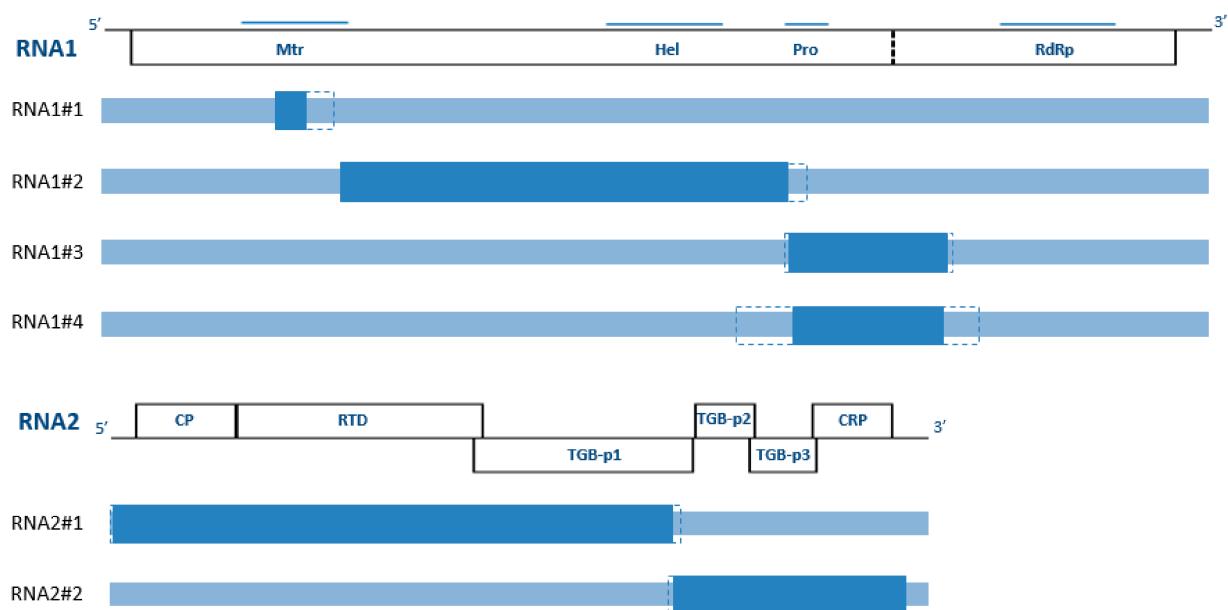
Figure 3. Recombination in RSNV genome. Genetic organization of the two RNA segments is indicated as in Figure 1. Dark blue rectangles indicate the location of the recombination events using RDP v4.974 [48] and dotted lines indicate the 99% of confident intervals. Five recombination events have been identified in this study.

Table 2. Recombination events detected in RSNV complete genome.

Recomb. Event [a]	Recombinant [a]	Major Parent	Minor Parent	Beginning Breakpoint [b]	Ending Breakpoint	Positive Methods [c]	p Value
RNA1 #1	M1-1	BF01Bama	Cor18	1026 (nt)	1213 (1129–1375)	G,B,M,C,T	1.24×10^{-5}
RNA1 #2	SL254	SF2.1	unknown	1392 (nt)	4097 (2818–4208)	R,M,C,S,T	4.49×10^{-5}
RNA1 #3	col	unknown	BF02Bama	4109 (4081-4148)	5063 (4999–5093)	R,G,B,M,C,S,T	3.42×10^{-37}
RNA1 #4	Be2	M1-1	Cor18	4121 (3787–4208)	5031 (4826–5228)	G,B,M,C,S	4.67×10^{-4}
RNA2 #1	Be2	unknown	BF02Bama	16 (1–354)	3360 (3337–3410)	R,M,S,T	4.18×10^{-13}
RNA2 #2 *	BR-LE01-19	unknown	BF02Bama	3360 (3322–3586)	4751 (nt)	R,G,B,M,C,S,T	3.88×10^{-12}

[a] as labeled in Figure 3, [b] the 99% confident intervals between brackets, nt: not determined, [c] algorithms implemented in RDP v4.974 software [48]: R, RDP, G, GeneConv, B, Bootscan, M, MaxChi, C, Chimaera, S, Siscan, T, 3Seq, * the same putative recombinant event was found in two sequences: BR-LE01-19 and col.

Phylogenetic analyses were performed without the recombinant domains (Figure 4). Two main lineages were detected, both with RNA1 and RNA2: one African lineage with isolates from Mali and Burkina Faso, and one South American lineage with isolates from Argentina and Brazil. This result contrasted with BNYVV for which no geographical structure has been detected between the previously reported types A and B [27] based on the phylogenetic trees with RNA1 and RNA2 (data not shown). The RSNV Colombian isolate, used as a reference until now, and the isolate SL254, showed a high divergence compared to the other isolates, suggesting the existence of other lineages. As described for BNYVV RNA3, 4 and 5 segments [28], we detected a putative reassortment event between RSNV RNA1 and RNA2. Indeed, RNA1 of the Brazilian isolate BR-LE01-19 belonged to the South American lineage, whereas its RNA2 was closer to the Colombian one. In addition, we noticed that although the RNA2 sequences of the Brazilian isolates BR-TT01-19 and

UFT/2019 contained the same deletion in TGBp1, such as isolates BR-LE01-19 and col, they still clustered in the South American lineage.
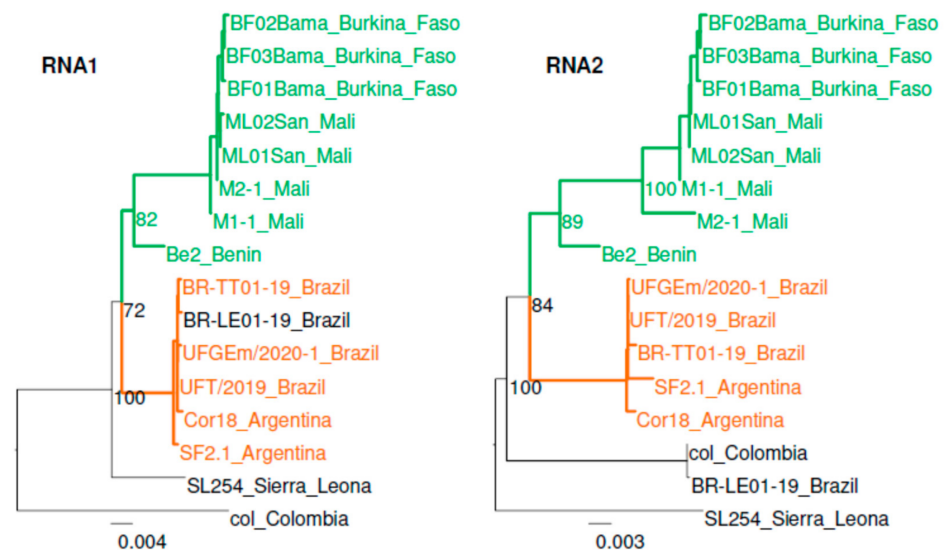

Figure 4. Phylogenetic relationships between RSNV genomic sequences. Maximum-likelihood (ML) phylogenetic trees were reconstructed with the sequences of the segments RNA1 and RNA2 of 16 RSNV isolates and without the recombinant regions. The trees were reconstructed using the PHYML-3.1 algorithm implemented in SEAVIEW v4.7 [39] under the best-fitted nucleotide substitution model (TN93+G) with gap treatment by pairwise deletion and with 100 bootstrap replications. The sequences clustering in the African and south African lineages are indicted in green and brown, respectively.

Lineage-specific amino-acids were investigated using DNAsp software (Figure 1, Tables S2 and S3). Twenty-six specific residues were identified in each RNA. In RNA1, non-synonymous parsimony informative sites were located mainly in the Mtr or Pro regions, or between Pro and RdRp (9, 5 and 5 mutations, respectively). Most identified residues were specific to the African lineage, except for isolates M1-1 and Be2. In RNA2, non-synonymous parsimony informative sites were located mainly in CP-RTD and TGBp1 and, to a lesser extent, in TGBp3 and CRP (16, 5, 3 and 2, respectively). However, only three African lineage specific sites were identified at position 346 in CP-RTD, at position 107 in TGBp1 and at position 79 in CRP, with the exception of the isolate Be2.

4. Discussion

The genetic diversity data in the genus *Benyvirus* were mainly restricted to the smallest RNA segments of BNYVV, additionally to its cp gene. Until recently, only four whole genomes of RSNV from South America have been available in the public databases, regardless of its wide distribution in Africa. In this study, we used a combination of three sequencing techniques to obtain the first corpus of African genomic sequences of RSNV. Genomic data thus increased to 16 sequences, including two isolates from Argentina. Firstly, we revealed a new feature in RSNV RNA2. All the new sequences and the sequence UFGEm/2020-1 showed a conserved but longer TGBp1 compared both to the other benyvirus sequences, and to the other RSNV sequences from Colombia and Brazil. The 81 amino acids inserted into the N-terminal part of the protein would result in intrinsically disordered (ID) domains (Figure S2). ID domains are frequent in viral proteins and can be involved in RNA interaction [49,50]. Intrinsic disorder has been previously reported in the N-terminal part of the barley stripe mosaic virus TGBp1, a hordeivirus [49]. Interestingly, RSNV and Hordei-like TGBp1 share another common feature: the presence of putative NoLS and NLS, both located in the RSNV insertion (Figure 2). Similar to animal viruses, NoLS and nucleolus have been pointed out as playing a role in several plant-virus

interactions [51]. Recently, Li et al. [52] have shown that the barley stripe mosaic virus TGBp1 hijacks the nucleolar protein Fibrillarin (Fib2) for cell-to-cell movement of the virus. Fib2 has also been shown to play a role in the long-distance movement of the rice stripe virus [53]. Further investigations should be performed to evaluate the functional role of this insertion in TGBp1. Notably, the absence of 81 residue-long insertions in the RSNV reference master genome from Colombia results in truncated sequences due to unmapped reads from the Illumina sequencing. The availability of new genomic data should now allow a reanalysis of the raw data, and to check the *tgbp1* gene length in the other sequences from Brazil. In addition, even if the isolate col has been sequenced using the Sanger method, we would suggest re-sequencing to confirm the absence of the insertion in the *tgbp1* gene. The collection and sequencing of other contemporaneous and/or sympatric isolates in Colombia would also be very interesting in order to decipher the evolutionary pathway of the TGBp1 indel. We noticed that the reference isolate col is one of the most divergent of the other isolates, even in RNA1, or outside TGBp1 in RNA2 (Figure 4).

The genetic diversity of the RNA1 and RNA2 showed contrasted characteristics depending on the benyvirus studied. Surprisingly, RNSV RNA1 showed a diversity index three times higher than expected, based on BNYVV data [29]. We detected significant positive Tajima's D values in the Mtr region, and in the genetic domain between Pro and RdRp, suggesting a balancing selection in RSNV RNA1 only. Clusters of RSNV-specific polymorphisms were identified in these domains. In addition to a comparison of RSNV/BNYVV, further investigations of the diversity of BNYVV housekeeping genes should be performed, to better understand its evolution. In our study, the low nucleotidic diversity, the detection of putative recombination events, and the limited number of RSNV sequences, still make reassortment detection difficult. Nevertheless, the phylogeny obtained with the two RNA segments of the Brazilian isolate BR-LE01-19, suggests a reassortment event as previously reported for BNYVV RNA3, 4 and 5 segments [28] or as shown between BNYVV and Beet soil-borne mosaic virus RNA1 and RNA2 [21]. By using some of our sequences publicly available, de Souza et al. (2021) have also proposed such a reassortment [54]. Such a possibility remains to be carefully verified, since at this stage one cannot rule out the possibility of simultaneous co-occurrences. As genetic reassortment potentially results in fitness advantages or disadvantages to the progeny virus [55], the viral life traits of this Brazilian RSNV isolate should also be compared to non-reassortant ones.

In the genus *Benyvirus*, inter-species recombination has been frequently discussed in connection with the co-occurrence of BNYVV and BSBMV [28]. However, the detection of intra-species recombination events has been poorly investigated, probably due to the lack of whole genome sequences, the low intra-species genetic diversity and the focus on RNA3 diversity. In addition, frequency of recombination appears to be much lower in segmented compared with multipartite viruses [55]. In this study, by increasing the amount of genetic data, we revealed the first recombination events in the RSNV genome. As expected, the inferred phylogenetic relationships of putative recombinant isolates were different, with or without recombinant domains (Figure S4). It is worth noting that the RNA1 phylogeny of the putative recombinant isolate M1-1 was maintained, probably due to the smaller size of the recombinant domain. In addition, we performed complementary phylogenetic analyses with partial publicly-available sequences of other isolates. Phylogenetic trees inferred from the helicase region in RNA1 and the *cp* gene in RNA2, revealed new putative variants in Burkina Faso (isolates BF-K1 and BF-VK1) and Mali (isolates MALI-B1, MALI-I1 and MALI-SK1) (Figure S5). Validation should be performed by the entire sequencing of these isolates and by recombination analyses. In conclusion, this study allowed (i) the identification of RSNV domains showing molecular evolution signatures with putative functional impacts and (ii) the definition of two main RSNV lineages emerging in America and re-emerging in Africa, and several putative recombinant isolates. New sequence data on benyviruses will be required to identify parental lineages, and to reconstruct the evolutionary history of the two RNA segments, as an aspect which appears more complex than previously thought.

References

1. Muthayya, S.; Sugimoto, J.D.; Montgomery, S.; Maberly, G.F. An overview of global rice production, supply, trade, and consumption. *Ann. N. Y. Acad. Sci.* **2014**, *1324*, 7–14. [CrossRef]
2. Fauquet, C.M.; Thouvenel, J.C. Association d'un nouveau virus en bâtonnet avec la maladie de la nécrose à rayures du riz en Côte-d'Ivoire. *C. R. Acad. Sci. Sér.* **1983**, *296*, 575–578.
3. Sereme, D.; Neya, B.J.; Bangratz, M.; Brugidou, C.; Quedraogo, I. First Report of Rice stripe necrosis virus infecting rice in Burkina Faso. *Plant Dis.* **2014**, *98*, 1451. [CrossRef]
4. Oludare, A.; Sow, M.; Afolabi, O.; Pinel-Galzi, A.; Hébrard, E.; Silué, D. First report of rice stripe necrosis virus infecting rice in Benin. *Plant Dis.* **2015**, *99*, 735. [CrossRef]
5. Decroës, A.; Bagayoko, I.; Mahillon, M.; Verhaegen, H.; Liénard, C.; Legrève, A.; Bragard, C. Detection of the Rice stripe necrosis virus causing rice crinkle disease and its vector Polymyxa graminis f. sp. colombiana in Mali. *Plant Dis.* **2017**, *12*, 2155. [CrossRef]
6. Tucker, M.J.; Celli, M.G.; Conteh, A.B.; Taylor, D.R.; Hébrard, E.; Poulicard, N. First report of Rice stripe necrosis virus infecting rice in Sierra Leone. *New Dis. Rep.* **2020**, *41*, 10. [CrossRef]
7. Morales, F.; Ward, E.; Castaño, M.; Arroyave, J.; Lozano, I.; Adams, M. Emergence and partial characterization of Rice stripe necrosis virus and its fungus vector in South America. *Eur. J. Plant Pathol.* **1999**, *105*, 643–650. [CrossRef]
8. Maciel, J.L.N.; de Moraes, M.G.; Almança, M.A.K.; Matsumura, A.T.S.; Falcade, J.H. Ocorrência do Vírus Rice stripe necrosis virus em Lavouras de Arroz do Rio Grande do Sul. *Fitopatol. Bras.* **2006**, *31*, 4115. [CrossRef]
9. Lozano, I.; Morales, F. Molecular characterisation of Rice stripe necrosis virus as a new species of the genus Benyvirus. *Eur. J. Plant Pathol.* **2009**, *124*, 673–680. [CrossRef]
10. Paz, L.; Espinoza, A.; Amano, Y. *El virus del "Entorchamiento" del arroz en Ecuador*; Boletín Divulgato no. 363; INIAP, Estación Experimental Litoral Sur: Guayaquil, Ecuador, 2009; Available online: http://repositorio.iniap.gob.ec/handle/41000/3830 (accessed on 19 March 2021).

11. Maurino, M.A.; Giménez-Pecci, M.; Kruger, R.D.; Agueda-Cúndom, M.; Gutierrez, S.A.; Celli, M.G. First report of Rice stripe necrosis virus in Argentina. *Crop Prot.* **2018**, *114*, 143–147. [CrossRef]

12. Gilmer, D.; Ratti, C.; ICTV Report Consortium. ICTV virus taxonomy profile: Benyviridae. *J. Gen. Virol.* **2017**, *98*, 1571–1572. [CrossRef]

13. McGrann, G.R.D.; Grimmer, M.K.; Mutasa-Göttgens, E.S.; Stevens, M. Progress towards the understanding and control of sugar beet rhizomania disease. *Mol. Plant Pathol.* **2009**, *10*, 129–141. [CrossRef]

14. Heidel, G.; Rush, C.; Kendall, T.L.; Lommel, S.A.; French, R. Characteristics of beet soilborne mosaic virus, a furo-like virus infecting sugar beet. *Plant Dis.* **1997**, *81*, 1070–1076. [CrossRef]

15. Lee, L.; Telford, E.B.; Batten, J.S.; Scholthof, K.B.; Rush, C.M. Complete nucleotide sequence and genome organization of Beet soilborne mosaic virus, a proposed member of the genus Benyvirus. *Arch. Virol.* **2001**, *146*, 2443–2453. [CrossRef]

16. Kondo, H.; Hirano, S.; Chiba, S.; Andika, I.B.; Hirai, M.; Maeda, T.; Tamada, T. Characterization of burdock mottle virus, a novel member of the genus Benyvirus, and the identification of benyvirus-related sequences in the plant and insect genomes. *Virus Res.* **2013**, *177*, 75–86. [CrossRef]

17. Tamada, T.; Kondo, H. Biological and genetic diversity of plasmodiophorid-transmitted viruses and their vectors. *J. Gen. Plant Pathol.* **2013**, *79*, 307–320. [CrossRef]

18. Erhardt, M.; Vetter, G.; Gilmer, D.; Bouzoubaa, S.; Richards, K.; Jonard, G.; Guilley, H. Subcellular localization of the Triple Gene Block movement proteins of Beet necrotic yellow vein virus by electron microscopy. *Virology* **2005**, *340*, 155–166. [CrossRef]

19. Zhang, L.; Wang, Z.; Wang, X.; Li, D.; Han, C.; Zhai, Y.; Yu, J. Two virus-encoded RNA silencing suppressors, P14 of Beet necrotic yellow vein virus and S6 of Rice black streak dwarf virus. *Chin. Sci. Bull.* **2005**, *50*, 305. [CrossRef]

20. Rahim, M.D.; Andika, I.B.; Han, C.; Kondo, H.; Tamada, T. RNA4-encoded p31 of beet necrotic yellow vein virus is involved in efficient vector transmission, symptom severity and silencing suppression in Roots. *J. Gen. Virol.* **2007**, *88*, 1611–1619. [CrossRef]

21. Laufer, M.; Mohammad, H.; Maiss, E.; Richert-Pöggeler, K.; Dall'Ara, M.; Ratti, C.; Gilmer, D.; Liebe, S.; Varrelmann, M. Biological properties of Beet soil-borne mosaic virus and Beet necrotic yellow vein virus cDNA clones produced by isothermal in vitro recombination: Insights for reassortant appearance. *Virology* **2018**, *518*, 25–33. [CrossRef] [PubMed]

22. Jupin, I.; Guilley, H.; Richards, K.E.; Jonard, G. Two proteins encoded by beet necrotic yellow vein virus RNA 3 influence symptom phenotype on leaves. *EMBO J.* **1992**, *11*, 479–488. [CrossRef] [PubMed]

23. Tamada, T.; Uchino, H.; Kusume, T.; Saito, M. RNA 3 deletion mutants of beet necrotic yellow vein virus do not cause rhizomania disease in sugar beets. *Phytopathology* **1999**, *89*, 1000–1006. [CrossRef] [PubMed]

24. D'Alonzo, M.; Delbianco, A.; Lanzoni, C.; Rubies Autonell, C.; Gilmer, D.; Ratti, C. Beet soilborne mosaic virus RNA4 encodes a 32 kDa protein involved in symptom expression and in virus transmission through Polymyxa betae. *Virology* **2012**, *423*, 187–194. [CrossRef]

25. Heijbroek, W.; Musters, P.M.S.; Schoone, A.H.L. Variation in pathogenicity and multiplication of Beet necrotic yellow vein virus (BNYVV) in relation to the resistance of sugar-beet cultivars. *Eur. J. Plant Pathol.* **1999**, *105*, 397–405. [CrossRef]

26. Simon-Loriere, E.; Holmes, E.C. Gene duplication is infrequent in the recent evolutionary history of RNA viruses. *Mol. Biol. Evol.* **2013**, *30*, 1263–1269. [CrossRef]

27. Ratti, C.; Clover, G.R.G.; Autonella, C.R.; Harjub, V.A.; Henry, C.M. A multiplex RT-PCR assay capable of distinguishing beet necrotic yellow vein virus types A and B. *J. Virol. Methods* **2005**, *124*, 41–47. [CrossRef]

28. Chiba, S.; Kondo, H.; Miyanishi, M.; Andika, I.B.; Han, C.; Tamada, T. The evolutionary history of Beet necrotic yellow vein virus deduced from genetic variation, geographical origin and spread and the breaking of host resistance. *Mol. Plant Microbe Int.* **2011**, *24*, 207–218. [CrossRef]

29. Zhuo, N.; Jiang, N.; Zhang, C.; Zhang, Z.Y.; Zhang, G.Z.; Han, C.G.; Wang, Y. Genetic diversity and population structure of beet necrotic yellow vein virus in China. *Virus Res.* **2015**, *205*, 54–62. [CrossRef]

30. Koenig, R.; Loss, S.; Specht, J.; Varrelmann, M.; Lüddecke, P.; Deml, G. A single U/C nucleotide substitution changing alanine to valine in the beet necrotic yellow vein virus P25 protein promotes increased virus accumulation in roots of mechanically inoculated, partially resistant sugar beet seedlings. *J. Gen. Virol.* **2009**, *90*, 759–763. [CrossRef]

31. Mariac, C.; Scarcelli, N.; Pouzadou, J.; Barnaud, A.; Billot, C.; Faye, A.; Kougbeadjo, A.; Maillol, V.; Martin, G.; Sabot, F.; et al. Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Mol. Ecol. Resour.* **2014**, *14*, 1103–1113. [CrossRef]

32. Afgan, E.; Baker, D.; van den Beek, M.; Blankenberg, D.; Bouvier, D.; Čech, M.; Chilton, J.; Clements, D.; Coraor, N.; Eberhard, C.; et al. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **2016**, *44*, W3–W10. [CrossRef]

33. Giardine, B.; Riemer, C.; Hardison, R.C.; Burhans, R.; Elnitski, L.; Shah, P.; Zhang, Y.; Blankenberg, D.; Albert, I.; Taylor, J.; et al. Galaxy: A platform for interactive large-scale genome analysis. *Genome Res.* **2005**, *15*, 1451–1455. [CrossRef] [PubMed]

34. Hébrard, E.; Pinel-Galzi, A.; Aderonke, O.; Poulicard, N.; Aribi, J.; Fabre, S.; Issaka, S.; Mariac, C.; Dereeper, A.; Albar, L.; et al. Identification of a Hypervirulent Pathotype of Rice Yellow Mottle Virus: A Threat to Genetic Resistance Deployment in West-Central Africa. *Phytopathology* **2018**, *108*, 299–307. [CrossRef]

35. Candresse, T.; Filloux, D.; Muhire, B.; Julian, C.; Galzi, S.; Fort, G.; Bernardo, P.; Daugrois, J.H.; Fernandez, E.; Martin, D.P.; et al. Appearances can be deceptive: Revealing a hidden viral infection with deep sequencing in a plant quarantine context. *PLoS ONE* **2014**, *9*, e102945. [CrossRef]

36. Filloux, D.; Murrell, S.; Koohapitagtam, M.; Golden, M.; Julian, C.; Galzi, S.; Uzest, M.; Rodier-Goud, M.; D'Hont, A.; Vernerey, M.S.; et al. The genomes of many yam species contain transcriptionally active endogenous geminiviral sequences that may be functionally expressed. *Virus Evol.* **2015**, *1*, vev002. [CrossRef]

37. Hall, T.A. Bioedit: A user-friendly biologically sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98. [CrossRef]

38. GenBank, National Center for Biotechnology Information—NCBI. Available online: http://www.ncbi.nlm.nih.gov (accessed on 9 April 2021).

39. Gouy, M.; Guindon, S.; Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **2010**, *27*, 221–224. [CrossRef] [PubMed]

40. Gibbs, A.J.; Torronen, M.; Mackenzie, A.M.; Wood, J.T.; Armstrong, J.S.; Kondo, H.; Tamada, T.; Keese, P.L. The enigmatic genome of Chara australis virus. *J. Gen. Virol.* **2011**, *92*, 2679–2690. [CrossRef] [PubMed]

41. Gaboriaud, C.; Bissery, V.; Benchetrit, T.; Mornon, J.P. Hydrophobic cluster analysis: An efficient new way to compare and analyse amino acid sequences. *FEBS Lett.* **1987**, *224*, 149–155. [CrossRef]

42. Romero, P.; Obradovic, Z.; Dunker, A.K. Sequence data analysis for long disordered regions prediction in the calcineurin family. *Genome Inform.* **1997**, *8*, 110–124.

43. Li, X.; Romero, P.; Rani, M.; Dunker, A.K.; Obradovic, Z. Predicting protein disorder for N-, C-, and internal regions. *Genome Inform.* **1999**, *10*, 30–40.

44. Scott, M.S.; Troshin, P.V.; Barton, G.V. NoD: A Nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC Bioinform.* **2011**, *12*, 317. [CrossRef] [PubMed]

45. Jiang, Z.; Zhang, K.; Li, Z.; Li, Z.; Yang, M.; Jin, X.; Cao, Q.; Wang, X.; Yue, N.; Li, D.; et al. The Barley stripe mosaic virus γb protein promotes viral cell-to-cell movement by enhancing ATPase-mediated assembly of ribonucleoprotein movement complexes. *PLoS Pathog.* **2020**, *16*, e1008709. [CrossRef]

46. Kosugi, S.; Hasebe, M.; Tomita, M.; Yanagawa, H. Systematic identification of yeast cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10171–10176. [CrossRef] [PubMed]

47. Librado, P.; Rozas, J. DNASP V5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]

48. Martin, D.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, vev003. [CrossRef]

49. Mishra, P.M.; Verma, N.C.; Rao, C.; Uversky, V.N.; Nandi, C.K. Intrinsically disordered proteins of viruses: Involvement in the mechanism of cell regulation and pathogenesis. *Prog. Mol. Biol. Transl. Sci.* **2020**, *174*, 1–78. [CrossRef]

50. Basu, S.; Bahadur, R.P. A structural perspective of RNA recognition by intrinsically disordered proteins. *Cell. Mol. Life Sci.* **2016**, *73*, 4075–4084. [CrossRef]

51. Taliansky, M.E.; Brown, J.W.; Rajamäki, M.L.; Valkonen, J.P.; Kalinina, N.O. Involvement of the plant nucleolus in virus and viroid infections: Parallels with animal pathosystems. *Adv. Virus Res.* **2010**, *77*, 119–158. [CrossRef]

52. Li, Z.; Zhang, Y.; Jiang, Z.; Jin, X.; Zhang, K.; Wang, X.; Han, C.; Yu, J.; Li, D. Hijacking of the nucleolar protein fibrillarin by TGB1 is required for cell-to-cell movement of Barley stripe mosaic virus. *Mol. Plant Pathol.* **2018**, *19*, 1222–1237. [CrossRef] [PubMed]

53. Zheng, L.; Du, Z.; Lin, C.; Mao, Q.; Wu, K.; Wu, J.; Wei, T.; Wu, Z.; Xie, L. Rice stripe tenuivirus p2 may recruit or manipulate nucleolar functions through an interaction with fibrillarin to promote virus systemic movement. *Mol. Plant Pathol.* **2015**, *16*, 921–930. [CrossRef] [PubMed]

54. de Souza, D.D.; de Queiroz, A.P.; Pereira, F.S.; de Campos Dianese, E.; Martins Fajardo, A.N.; Lau, D.; da Silva, L.A.; Morais Ribeiro, B.; Guedes Coehlo, A.S.; de Souza Aguiar, R.W.; et al. Molecular characterization and sequence analysis of four Brazilian rice stripe necrosis virus isolates. *Arch. Virol.* **2021**, 1–5. [CrossRef]

55. McDonald, S.M.; Nelson, M.I.; Turner, P.E.; Patton, J.T. Reassortment in segmented RNA viruses: Mechanisms and outcomes. *Nat. Rev. Microbiol.* **2016**, *14*, 448–460. [CrossRef]