Virus Research 259 (2019) 1–9

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Evolutionary origin of 2A-like sequences in *Totiviridae* genomes

Juliana G.S. de Lima[a,c], Diego G. Teixeira[b,c], Tiago T. Freitas[a,d], João P.M.S. Lima[b,c], Daniel C.F. Lanza[a,c,d,*]

[a] Applied Molecular Biology Lab - LAPLIC, Department of Biochemistry, Federal University of Rio Grande do Norte, Natal, RN, Brazil
[b] Laboratory of Metabolic Systems and Bioinformatics - LASIS, Department of Biochemistry, Federal University of Rio Grande do Norte, Natal, RN, Brazil
[c] Postgraduate Program in Biochemistry, Federal University of Rio Grande do Norte, Natal, RN, Brazil
[d] Postgraduate Program in Science, Technology and Innovation, Federal University of Rio Grande do Norte, Natal, RN, Brazil

ARTICLE INFO

Keywords:
Pseudo 2A-site
Artivirus
Giardiavirus
2A-like ranking
Arthropod totivirus
IMNV

ABSTRACT

In recent years there has been a significant increase in the number of new species potentially belonging to the *Totiviridae* family. Most of these new viruses have not yet been covered by the Committee on Taxonomy of Viruses (ICTV) official classification. In this study, a phylogenetic analysis including new sequences of *Totiviridae* candidates revealed a clade including *Giardiavirus* and a great diversity of new totiviruses, which infect arthropods, protozoa and mollusc. This expanded *Giardiavirus* clade comprises two monophyletic groups, one of them including *Giardia lamblia virus* (GLV) grouped with viruses that infect arthropods and vertebrates (GLV-like group), and the other includes the previously proposed Artivirus group (IMNV-like group). A screening of the members of the GLV-like group in search of genomic elements already described in IMNV-like group revealed the existence of sites with a high propensity to become 2 A-like oligopeptides, mainly in a specific subgroup of arthropod viruses, suggesting that these viruses preserved ancestral characteristics. The existence of these "pseudo 2 A-sites" associated to phylogenetic reconstruction indicates that these sequences appear at a decisive stage for viral evolution. If they are changed to functional 2 A-like sequences, an irreversible route to increase the genome complexity will be initiated.

1. Introduction

The *Totiviridae* family comprises non-enveloped icosahedral virions of ~40 nm, containing monopartite double stranded RNA (dsRNA) genomes. Their genomes range from 4.6 to 7.0 kbp in size, and mostly of them are organized in two Open Reading Frames (ORFs). The ORF1 encodes a capsid protein which generally present sizes between 70–100 kDa, and ORF2 encodes an RNA-dependent RNA polymerase (RdRp) (Wickner et al., 2012). According to the latest classification by the Committee on Taxonomy of Viruses (ICTV), the *Totiviridae* family comprises five genera: *Giardiavirus*, *Leishmaniavirus* and *Trichomonasvirus* comprising viruses that infect protozoa, and *Totivirus* and *Victorivirus* which exclusively infect fungi.

In recent years there has been a significant increase in the number of studies describing new species potentially belonging to *Totiviridae* family, and most of them have not yet been covered by official ICTV classification. At least 37 new *Totiviridae* genomes were added to the National Center for Biotechnology Information (NCBI) databank in the last 5 years. Surprisingly, some of these had an unexpected host distribution, including a variety of insects, some plants and even fish species (Koyama et al., 2015; Martinez et al., 2016; Chen et al., 2016; Mor and Phelps, 2016b). Due to this large host heterogeneity, the *Totiviridae* family has increasingly become an excellent model for studies that aim to clarify the relationships between phylogeny and structural genome features.

In fact, despite the small number of studies on the subject, there is evidence that certain genomic elements occur exclusively in specific groups within the *Totiviridae* family. In this context, one of the principal examples are the 2A-like sequences. 2A is an oligopeptide sequence with eight or nine amino acid signatures that mediate a ribosome 'skipping' effect, analogous to internal ribosome entry sequences (IRES), producing an apparent co-translation 'cleavage' of polyproteins without needing a proteinase (Donnelly et al., 2001a, 2001b; Luke et al., 2008). The existence of 2A-like sequences in the *Totiviridae* family was first reported in *Infectious myonecrosis virus* (IMNV), a virus that infects penaeid shrimps, by Nibert (2007). From there, the presence of 2 A-like sequences has been reported in other genomes such as *Armigeres subalbatus totivirus* (AsV), *Drosophila melanogaster totivirus* (DTV),

* Corresponding author at: Laboratório de Biologia Molecular Aplicada – LAPLIC, Departamento de Bioquímica, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, RN, CEP: 59072-970, Brazil.
E-mail address: danielclanza@gmail.com (D.C.F. Lanza).

https://doi.org/10.1016/j.virusres.2018.10.011
Received 31 May 2018; Received in revised form 28 August 2018; Accepted 15 October 2018
Available online 16 October 2018
0168-1702/

Omono river virus (OMRV), *Tianjin totivirus* (ToV) and *Golden shiner totivirus* (GSTV) (Zhai et al., 2010; Isawa et al., 2011; Yang et al., 2012; Mor and Phelps, 2016b).

These viruses (OMRV, ToV, DTV, AsV, IMNV and GSTV) present one or two 2A-like sequences and one cleavage site located upstream of the major capsid protein (MCP) coding region in their OFR1. Because of this, they potentially encode two or three proteins in addition to MCP. One of these is a RNA binding protein (RBP) with a double-stranded RNA-binding domain (DSRM), and the other is the small protein 2 (SP2), a hypothetical protein that may be a component of the IMNV capsid and can be related to virus cell entry and extracellular transmission (Dantas et al., 2016).

Zhai et al. (2010) suggested the creation of Artivirus as a new genus comprising arthropod *Totiviridae* species based on IMNV, DTV and AsV genome characteristics, phylogenetic relationships and host type. Further genome and phylogenetic analyses performed by Dantas et al. (2016) supported the existence of typical characteristics observed in IMNV, DTV, AsV, OMRV and ToV.

In this study we have updated the phylogenetic relations of *Totiviridae* family, including all newly described species so far, and propose the possible origin of the 2 A-like sequences in this family.

2. Methods

2.1. Phylogenetic analysis

The 90 protein sequences used for reconstruction of the RdRp-based phylogenetic tree were obtained from NCBI Taxonomy database (Supplementary material 1). *Totiviridae* sequences were selected according to the following criteria: (1) sequences annotated as "*Totiviridae*" or "unclassified *Totiviridae*" deposited before October/2017; (2) in cases where more than one lineage was available for the same species, only the larger size sequence was used; (3) sequences containing more than 300 amino acids. Although the sequences designated by ToV and OMRV possibly correspond to the same organism, these two sequences were maintained in the analysis since they correspond to different annotations. Sequences designated as "diatom colony associated dsRNA virus", "environmental samples" or those that were not described/characterized in published articles were not considered. Four sequences (ADO60933; YP_009209482; YP_003541123; CAJ34335) that reduce the posteriori values of their respective groups and one sequence (YP_009094186) that grouped with outgroup sequences in phylogenetic analysis were removed. RdRp sequences from *Amalgaviridae* family were used as the outgroup.

Sequence alignments were performed using the online version of MAFFT (Yamada et al., 2016), with structural alignment option guided by the RT_like family domain (cd00304) obtained from the NCBI *Conserved Domains Database* (CDD) (Marchler-Bauer et al., 2016). Afterwards, the alignment was submitted to the ProtTest program (Darriba et al., 2011) in order to estimate the best amino acid substitution model, followed by a Bayesian phylogenetic analysis using the BEAST v2.4 program (Bouckaert et al., 2014). The Bayesian inference was performed from two independent runs using the LG substitution model, a strict type molecular clock with a discrete trait informing which group each one of the viruses could belong to in a constant population coalescing model. The Markov chain Monte Carlo (MCMC) was performed with 10^7 generations, with log and tree samplings per 1000 generations. The consensus tree was obtained from the combination of both independent runs with a 10% burning in of the initial states. The phylogenetic inference quality was analyzed in the Tracer v1.6 program (Drummond et al., 2012) and the tree topology was generated and edited using the FigTree v1.4 program (Drummond et al., 2012).

2.2. Genome structure and ORF1 organization of the members of expanded Giardiavirus group

Genome schemes were elaborated based in NCBI sequences: NC_007915.3; NC_030295.1; NC_014609.1; NC_013499.1; NC_017084.1; AB555544.1; NC_003555.1; NC_015639.1; NC_025218.2; KX148550.1; NC_027212.1; NC_029312.1; and NC_029302.1, and also according to the data from previous descriptions (Koyama et al., 2015; Dantas et al., 2016; Fauver et al., 2016; Koyama et al., 2016; Martinez et al., 2016; Mor and Phelps, 2016a, b). The FoldIndex© web server tool (Prilusky et al., 2005) was used to predict folded or unfolded regions. For GSTV characterization, DSRM domain and putative SP2 sequences of GSTV were aligned using MUSCLE (Edgar, 2004) and T-Coffee (Notredame et al., 2000) with default parameters and manually adjusted using the Jalview v.2.8 interface (Waterhouse et al., 2009). 3D protein models were generated in I-TASSER (Zhang, 2008; Roy et al., 2010) and model scores were evaluated using MOLPROBITY (Chen et al., 2010) and SWISS-MODEL (Arnold et al., 2006).

2.3. Pseudo 2 A site analysis and cleavage site prediction

2 A-like sequences described in Zhai et al. (2010); Isawa et al. (2011); Yang et al. (2012) and Mor and Phelps (2016b) were aligned against *Leptopilinia boulard toti-like virus* (LbTV), *Anopheles totivirus* (AToV), *Australian Anopheles totivirus* (AATV), *Camponotus nipponicus virus* (CNV) and *Camponotus yamaokai virus* (CYV) ORF1 protein sequences (YP_009072447.2; KX148550.1; YP_009417300.1; YP_009230207.1; YP_009143312.1) using BlastP algorithm and Geneious 7.1.3, and colored by its polarity and hydrophobicity. 2 A-like sequences from IMNV-like group were aligned with the dsRNA bank of NCBI using BlastP algorithm to search for pseudo 2 A site sequences in other dsRNA viruses. The sequences that showed at least 63% of identity were compared with 2 A-like sequences experimentally validated and notified in the literature (Supplementary material 2) using the 2 A-like ranking software. The codon weights were established based on the proportion of invariant sites for each codon position using Data Analysis in Molecular Biology and Evolution (DAMBE) software (Xia and Xie, 2001).

2.4. 2A-like ranking software

The 2A-like ranking software was developed to quantify the propensity of a particular sequence in becoming a functional 2 A-like by means of random mutations. The algorithm works in five steps: (1) The potential codons that encode that amino acid are identified; (2) each nucleotide position is analyzed per codon. The matches are scored according to the pre-defined weights, assigned by the user for each nucleotide position within the codon. By default, the weights for the first codon position is 0.3, the second is 0.6 and the third is 0.1; (3) the initial score is calculated by summing the individual scores of each codon; (4) the weight of each amino acid of the query sequence is defined, in other words, how much each codon can score depending on the sequence size. For example, in a sequence of 9 nucleotides that corresponds to 3 amino acids, the weight of each codon (1 amino acid) is defined by the division of 1 by the number of codons (3 codons), which results in 0.33. (5) Finally, after the sum of the nucleotide scores, this value is multiplied by the value of each amino acid and then all the scores for each amino acid are summed. A more detailed example is shown in Supplementary material 3. The software code is available at https://github.com/LAPLIC/2A-like-ranking.

3. Results and discussion

3.1. Phylogenetic analysis

Phylogenetic analysis based on the RdRp sequences revealed the

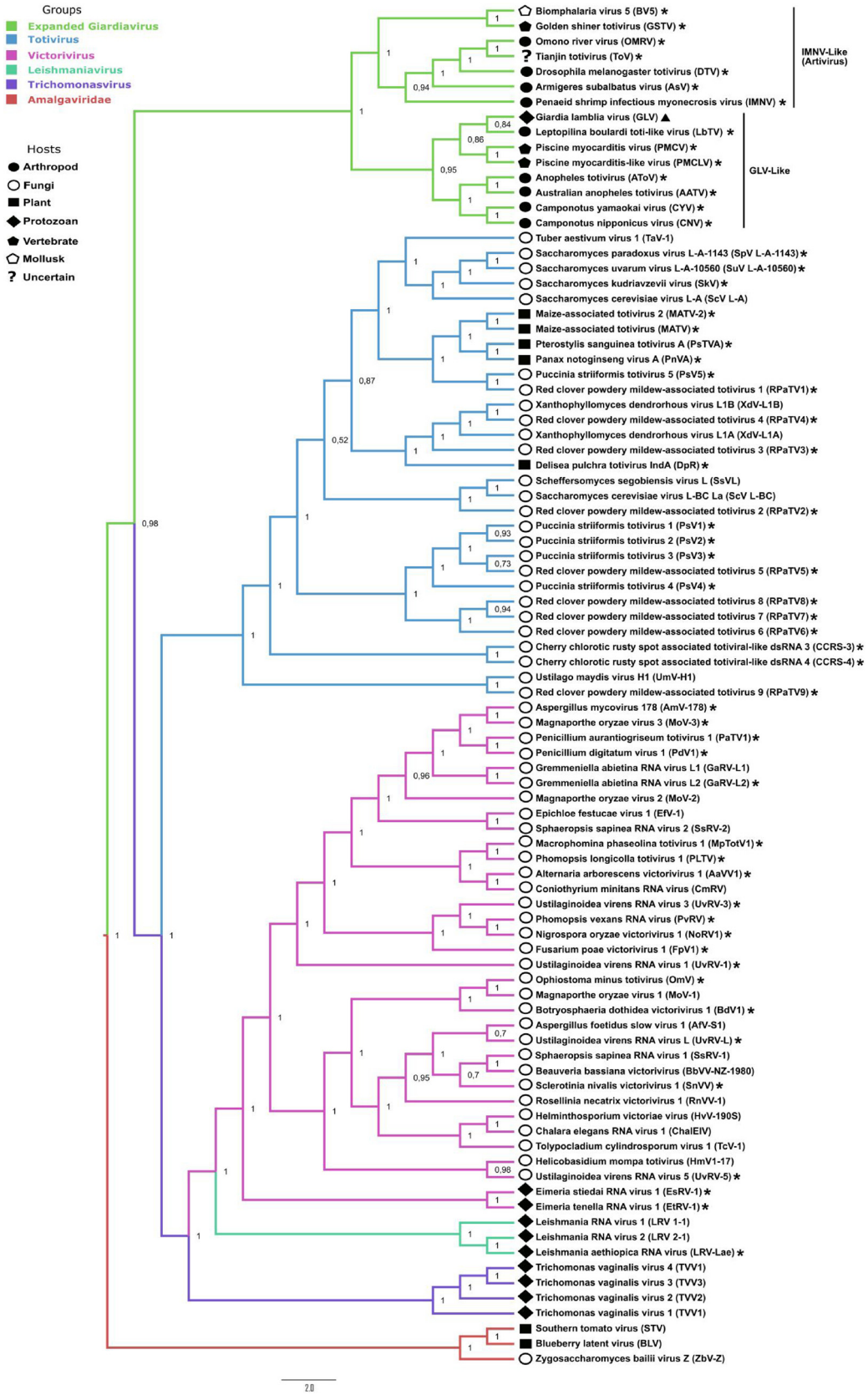

Fig. 1. Phylogenetic relationships between *Totiviridae* family members. The tree was calculated from an alignment of 90 RdRp amino acid sequences comprising representative members of the *Totiviridae* family and tentatively assigned viruses, using Bayesian inference. Sequences not formally assigned to ICTV taxonomy are highlighted with asterisks. Numbers in branch nodes indicate posterior probabilities. Colored branches represent the ancestry for each group. The black triangle indicates the single virus formally classified within the *Giardiavirus*. The tentatively groups GLV-like and IMNV-like (Artivirus) are highlighted. Different shapes in front of each name indicate the host of each respective virus.

existence of six monophyletic taxa in *Totiviridae* family (Fig. 1). These groups include the genera adopted by ICTV such as *Giardiavirus*, *Totivirus*, *Victorivirus*, *Leishmaniavirus* and *Trichomonasvirus*, as shown in the dendrogram presented in Fig.1.

Among the 90 species submitted to phylogenetic analysis, 59 unclassified viruses are grouped within genera formally classified by ICTV (Table 1). The highest number of new viruses (24 viruses) is clustered within the clade of the *Totivirus* genus. Interestingly, 5 among these 24

Table 1
Suggested classification for the totiviruses analyzed in this article.

Virus name	Abbreviation	First described in	Proposed genus
Leishmania aethiopica RNA virus	LRV-Lae	(Zangger et al., 2014)	*Leishmaniavirus*
Cherry chlorotic rusty spot associated totivirallike dsRNA 3	CCRS-3	(Kozlakidis et al., 2006)	*Totivirus*
Cherry chlorotic rusty spot associated totivirallike dsRNA 4	CCRS-4	(Kozlakidis et al., 2006)	
Delisea pulchra totivirus IndA	DpR	(Lachnit et al., 2016)	
Pterostylis sanguinea totivirus A	PsTVA	(Ong et al., 2018)	
Puccinia striiformis totivirus 1	PsV1	(Zheng et al., 2017)	
Puccinia striiformis totivirus 2	PsV2	(Zheng et al., 2017)	
Puccinia striiformis totivirus 3	PsV3	(Zheng et al., 2017)	
Puccinia striiformis totivirus 4	PsV4	(Zheng et al., 2017)	
Puccinia striiformis totivirus 5	PsV5	(Zheng et al., 2017)	
Red clover powdery mildew-associated totivirus 1	RPaTV1	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 2	RPaTV2	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 3	RPaTV3	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 4	RPaTV4	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 5	RPaTV5	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 6	RPaTV6	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 7	RPaTV7	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 8	RPaTV8	(Kondo et al., 2016)	
Red clover powdery mildew-associated totivirus 9	RPaTV9	(Kondo et al., 2016)	
Saccharomyces kudriavzevii virus	SkV	(Rowley et al., 2016)	
Saccharomyces paradoxus virus L-A-1143	SpV L-Al1143	(Rodriguez-Cousiño et al., 2017)	
Saccharomyces uvarum virus L-A-10560	SuV L-A-10560	(Rodriguez-Cousiño et al., 2017)	
Maize-associated totivirus	MATV	(Quito-Avila et al., 2016)	
Maize-associated totivirus 2	MATV-2	(Quito-Avila et al., 2016)	
Panax notoginseng virus A	PnVA	(Guo et al., 2016)	
Alternaria arborescens victorivirus 1	AaVV1	(Komatsu et al., 2016)	*Victorivirus*
Aspergillus mycovirus 178	AmV-178	(Hammond et al., 2008)	
Botryosphaeria dothidea victorivirus 1	BdV1	(Zhai et al., 2015)	
Eimeria stiedai RNA virus 1	EsRV-1	(Xin et al., 2016)	
Eimeria tenella RNA virus 1	EtRV-1	(Wu et al., 2016a)	
Fusarium poae victorivirus 1	FpV1	(Osaki et al., 2016)	
Gremmeniella abietina RNA virus L2	GaRV-L2	(Tuomivirta and Hantula, 2005)	
Macrophomina phaseolina totivirus 1	MpTotV1	(Marzano et al., 2016)	
Magnaporthe oryzae virus 3	MoV-3	(Tang et al., 2015)	
Nigrospora oryzae victorivirus 1	NoRV1	(Zhong et al., 2016)	
Ophiostoma minus totivirus	OmV	(Doherty et al., 2007)	
Penicillium aurantiogriseum totivirus 1	PaTV1	(Nerva et al., 2016)	
Penicilium digitatum virus 1	PdV1	(Niu et al., 2016)	
Phomopsis longicolla totivirus 1	PLTV	(Marzano et al., 2016)	
Phomopsis vexans RNA vírus	PvRV	(Zhang et al., 2015)	
Ustilaginoidea virens RNA virus 1	UvRV1	(Zhong et al., 2014a)	
Ustilaginoidea virens RNA virus 3	UvRV-3	(Zhong et al., 2014b)	
Ustilaginoidea virens RNA virus 5	UvRV-5	(Zhong et al., 2017)	
Ustilaginoidea virens RNA virus L	UvRV-L	(Jiang et al., 2015)	
Sclerotinia nivalis victorivirus 1	SnVV	(Wu et al., 2016b)	
Anopheles totivirus	AToV	(Fauver et al., 2016)	*Expanded*
Armigeres subalbatus virus	AsV	(Zhai et al., 2010)	*Giardiavirus*
Australian anopheles totivirus	AATV	(Colmant et al., 2017)	
Biomphalaria virus 5	BV5	(Galinier et al., 2017)	
Camponotus yamaokai virus	CYV	(Koyama et al., 2015)	
Camponotus nipponicus virus	CNV	(Koyama et al., 2016)	
Drosophila melanogaster totivirus	DTV	(Wu et al., 2010)	
Golden shiner totivirus	GSTV	(Mor and Phelps, 2016b)	
Leptopilinia boulardi toti-like virus	LbTV	(Martinez et al., 2016)	
Omono river virus	OMRV	(Isawa et al., 2011)	
Penaeid shrimp infectious myonecrosis virus	IMNV	(Poulos et al., 2006)	
Piscine myocarditis virus	PMCV	(Haugland et al., 2011)	
Piscine myocarditis-like virus	PMCLV	(Mor and Phelps, 2016a)	
Tianjin totivirus	ToV	(Yang et al., 2012)	

viruses (*Delisea pulchra totivirus IndA* (DpR), *Maize-associated totivirus* (MATV), *Maize-associated totivirus 2* (MATV-2), *Panax notoginseng virus A* (PnVA), *Pterostylis sanguinea totivirus A* (PsTVA)) do not infect fungi, but plants. The *Victorivirus* group received the second highest number of individuals, totaling 20 new viruses. Also in this group, phylogenetic analysis not only revealed viruses that infect fungi, but also 2 viruses that infect protozoa (*Eimeria stiedai RNA virus 1* (EsRV-1) and *Eimeria tenella RNA virus 1* (EtRV-1)). New viruses were also observed in *Leishmaniavirus* (1 virus) and *Giardiavirus* clades (14 viruses). Recently discovered viruses that infect the *Biomphalaria glabrata* mollusc or the *Notemigonus crysoleucas* fish, called BV5 and GSTV respectively, were grouped in the proposed Artivirus group, in addition to the other previously described arthropod totiviruses. Interestingly, the phylogenetic analysis revealed that five recently described arthropod viruses (LbTV, AToV, AATV, CYV and CNV) and two vertebrate viruses (PCMV and PMCLV) belong to the same clade of *Giardia lamblia virus* (Fig. 1).

These results suggested a new interpretation on the classification of the groups within the *Totiviridae* family. If so far there was a coincidence between viruses that infect hosts belonging to a unique kingdom inside each group, from now on the host diversity inside groups introduces that the evolutionary mechanisms cannot be restricted by a host type. In this context, the classification of *Giardiavirus*

as a protozoan group comprising only one species could be reviewed, considering the closely phylogenetic relationship of *Giardia lamblia virus* with some arthropod and vertebrate viruses (Fig. 1). The proposition that the Artivirus genus only contains viruses that infect arthropods could also be revised, taking into account the presence of viruses that infect vertebrates and mollusc in the same group.

Therefore, the two monophyletic groups observed within the expanded *Giardiavirus* clade will be herein referred to as IMNV-like and GLV-like (in reference to the first described virus of each group), to avoid any misinterpretation with groups formally accepted by ICTV. In this way, it would be interesting to define other characteristics that would allow a more robust criterion to classify the viruses within the expanded *Giardiavirus* clade.

3.2. Genome organization and ORF1 characteristics of the expanded *Giardiavirus* group

Since phylogenetic analyses were restricted to RdRp, the next step was to search characteristics related to the structural organization of ORF1 that would clarify the classification of the expanded *Giardiavirus* clade. The analysis of ORF1 organization confirmed that viruses of the IMNV-like group have unique characteristics that could be better defined, such as the existence of 2 A-like sequences, putative proteins RBP and SP2 and a cleavage site between SP2 and MCP, as illustrated in Fig. 2A. It was also possible to observe some conservation in the folding profile of these elements encoded by ORF1 according to the folding prediction shown in Fig. 2A below each scheme. Thus, 3D models were calculated as these elements had not yet been characterized for GSTV (Supplementary material 4 A). The DSRM model predicted two α-helices and a β-sheet, like the models previously described from the other IMNV-like viruses (Dantas et al., 2016). A putative SP2 located upstream of MCP was also identified, and the 3D model of this protein revealed a protein rich in α-helices (Supplementary material 4 A). The reliability scores for DSRM and SP2 3D models are detailed in Supplementary material 5. It was not possible to verify the existence of 2A-like sequences, or cleavage sites in BV5 since only the RdRp sequence was deposited in the NCBI database (Galinier et al., 2017).

There are no evident elements enabling to determine a consensus in ORF1 organization within the GLV-like group (Fig. 2B). Due to this, the ORF1 of these viruses were screened by means of sequence alignments, adopting the elements of IMNV-like viruses as probes (RBP, SP1, SP2, cleavage site and 2 A-like sequences) using the BlastP algorithm. Although it was not possible to observe evidence of RBP and SP2, it was notable to observe the existence of a probable cleavage site in LbTV ORF1, similar to that already described for the IMNV-like viruses (Fig. 3A). Another exciting finding was the identification of sequences

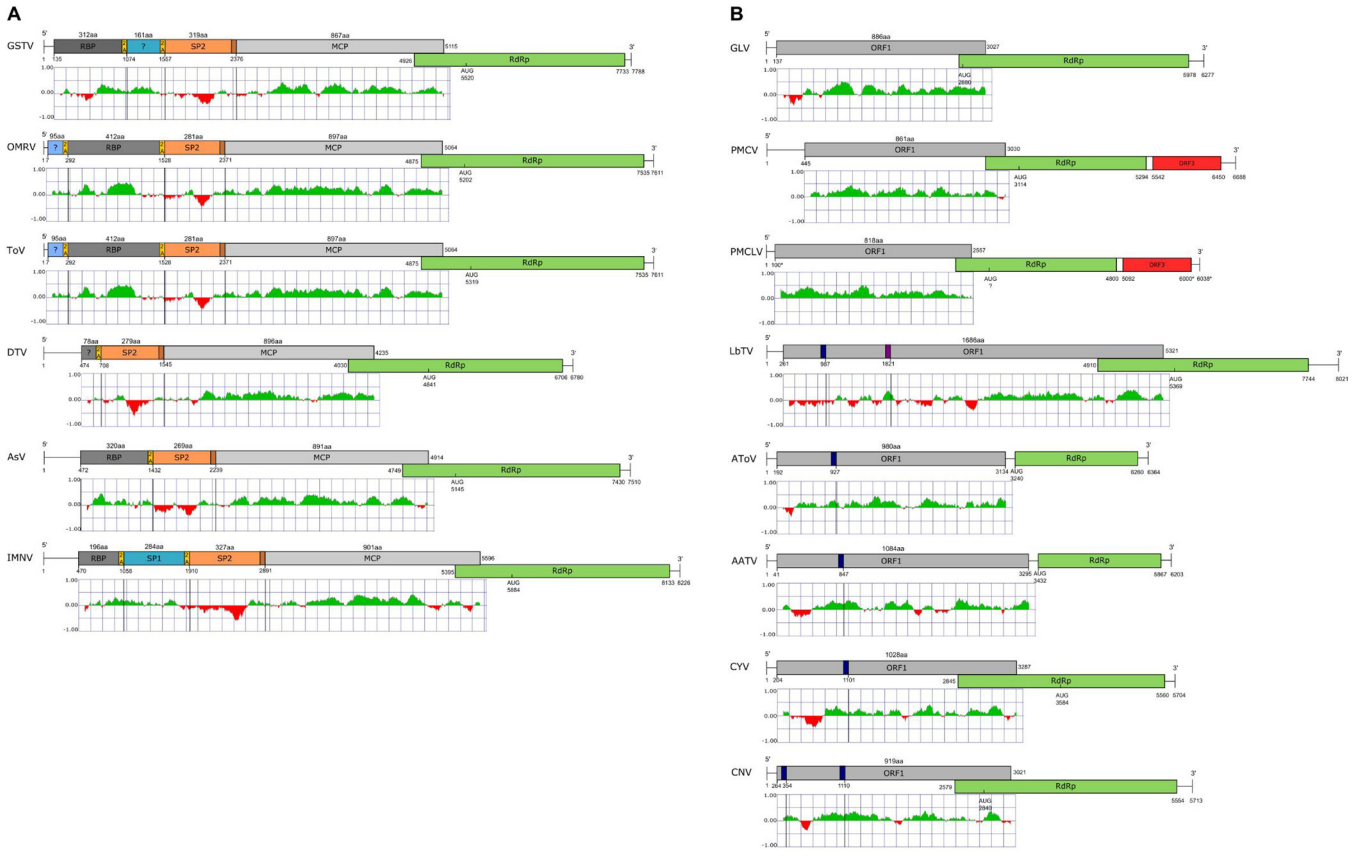

Fig. 2. Schematic presentation of ORF1 polypeptide. Schematic representations comparing full genomes of viruses grouped in the IMNV-like (A) and GLV-like (B) groups according to the phylogenetic analyses. The coding regions of the proteins are represented by rectangles of different colors: 2 A-like sequences (2 A) as yellow, RNA binding protein (RBP) as dark grey, small protein 1 (SP1) as light blue, small protein 2 (SP2) as orange, major capsid protein (MCP) as light grey and RNA-dependent RNA polymerase (RdRp) in green. Positions of the initial nucleotide referenced by the complete genome and the size of each predicted polypeptide encoded by ORF1 are represented by the numbers below and above the bars respectively. The brown and purple rectangles represent the putative cleavage sites. Dark blue rectangles represent the pseudo 2 A-site positions in GLV-like members. Graphs below each genome scheme indicate the folded (green) and unfolded (red) regions predicted by FoldIndex for ORF1, x-axis represent amino acid positions and the y-axis represents the probability scores. Asterisks indicate that 5′UTR and OFR3 were not completely sequenced and 3′UTR was not sequenced. Abbreviations: OMRV - GSTV - *Golden shiner totivirus*; *Omono river virus*; ToV - *Tianjin totivirus*; DTV - *Drosophila melanogaster totivirus*; AsV - *Armigeres subalbatus virus*; IMNV - *Infectious myonecrosis virus*; GLV - *Giardia lamblia virus*; PMCV - *Piscine myocarditis virus*; PMCLV - *Piscine myocarditis-like virus*; LbTV - *Leptopilina boulardi toti-like virus*; AToV - *Anopheles totivirus*; AATV - *Australian anopheles totivirus*; CYV - *Camponotus yamaokai virus*; CNV - *Camponotus nipponicus virus*.

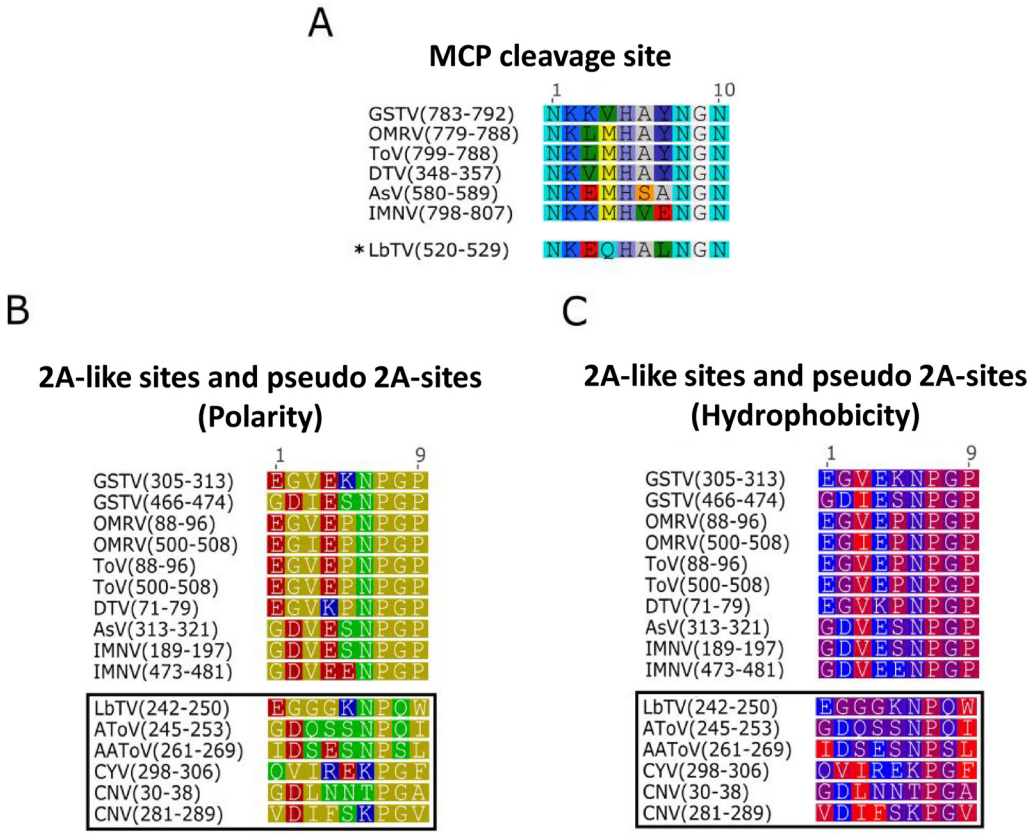

Fig. 3. Cleavage sites, 2A-like sequences and pseudo 2A-sites characterization. (A) Putative MCP cleavage sites identified in ORF1 sequences colored by amino acid type. Asterisks indicate the new putative cleavage site described in this article. (B, C) Amino acid sequences of 2A-like and pseudo 2A-sites identified in ORF1 polypeptide colored by polarity (B) and hydrophobicity (C). Numbers in parentheses correspond to the positions of each site in their respective ORF1 amino acid sequences. Sequences in the box correspond to pseudo 2 A-sites. Abbreviations: OMRV - GSTV - *Golden shiner totivirus*; *Omono river virus*; ToV - *Tianjin totivirus*; DTV - *Drosophila melanogaster totivirus*; AsV - *Armigeres subalbatus virus*; IMNV - *Infectious myonecrosis virus*; LbTV - *Leptopilina boulardi toti-like virus*; AToV - *Anopheles totivirus*; AATOV - *Australian anopheles totivirus*; CYV - *Camponotus yamaokai virus*; CNV - *Camponotus nipponicus virus*.

Table 2
Proportion of invariable sites for each codon position.

	1º codon position	2º codon position	3º codon position	1º and 2º codon position	Whole sequence
Artivirus	0,23	0,44	0	0,49713	0,25678
Giardiavirus	0,17170	0,02078	0	0,1111	0,02965

resembling 2 A-like sites in some GLV-like genomes (Figs. 2B and 3 B and C).

3.3. Pseudo 2A-site sequences

Screening the LbTV, AToV, AATV, CNV and CYV genomes using the previously described 2A-like sequences as probes revealed six sequences: LbTV(242-250) AToV(245-253), AATV(261-269), CNV(30-38), CNV(281-289) and CYV(298-306), which shared at least 50% identity with at least one of the original 2 A-like sequences. Indeed, the polarity and hydrophobicity profiles of these amino acid sequences resemble those observed in functional 2 A-like sequences, indicating that they may conserve partial characteristics (Fig. 3B and C); so these sequences were called "pseudo 2A-sites".

Pseudo 2A-sites probably do not produce a ribosome ‘skipping’ effect since some changes in specific amino acids prevents the cleavage of the polypeptide (Donnelly et al., 2001a, 2001b; Luke et al., 2008). However, these pseudo 2A-sites could be remnants of some common ancestor of the expanded *Giardiavirus* clade. In this way, it would be interesting to verify if the nucleotide sequences encoding these regions would have some propensity to generate a functional 2A-like, taking into account random mutation events.

3.4. 2A-like ranking software

We developed a software for determining the propensity of a given nucleotide sequence "to become" a 2 A-like coding sequence. This software is called “2A-like ranking” and quantifies the minimum number of substitutions that would be necessary in each nucleotide position of a query sequence, assigning points to each match (coincidence) between query and reference sequences (in this case the reference being a known 2A-like coding sequence). When the match occurs in the first two nucleotides of each codon the assigned score is higher, since the mutation effect in these positions is more pronounced, in agreement with the proportion of invariable sites in 2A-like sequences (Table 2). Finally, the score obtained by each codon is summed, and sequences with higher scores would be those with higher propensity to originate a functional 2A-like coding region. In other words, the software assigns the highest scores to sequences that need fewer mutation events to generate a 2A-like coding region.

3.5. 2A-like evolution in Totiviridae

The propensity of the identified pseudo 2A-sites to become functional 2A-like sequences was assessed using the 2A-like ranking software (Fig. 4A). In parallel, an analysis of similarity between the same sequences was performed using the BLOSUM62 matrix to encompass different scores for amino acid substitutions (Fig. 4 B). As expected, a high similarity with a functional 2 A-like sequence does not directly correlate with a greater propensity to become a functional 2 A-like by random mutation events. The results from BLOSUM45 and BLOSUM80 matrices produced similar results (data not shown). According to the 2 A-like ranking, the 2 A-site with the least propensity to become a functional 2 A-like belongs to LbTV. Interestingly, the other viruses (AToV, AATV, CYV, CNV) which present the most likely 2 A-sites to become functional form a specific monophyletic subgroup within the GLV-like group (Fig. 1).

A

	GDVEENPGP (IMNV)	GDVESNPGP (IMNV)	HDIETNPGP (EMCV)	GDVELNPGP (PoRV-C)	GDIELNPGP (BoRV-C)	GDIESNPGP (BmCPV-1)	HDVEMNPGP (TMCV)	GDVESNPGP (FMDV)
EGGGKNPQW (LbTV) GAAGGCGGGGGAAAGAA CCCGCAGTGG	0.500	0.467	0.444	0.444	0.400	0.422	0.489	0.467
GDQSSNPQI (AToV) GGTGACCARTCTTCAAAC CCTCAGATC	0.478	0.578	0.444	0.522	0.522	0.578	0.367	0.578
IDSESNPSL (AATV) ATAGACTCGGAATCCAAC CCTTCATTG	0.489	0.611	0.544	0.533	0.522	0.600	0.478	0.611
QVIREKPGF (CYV) CAAGTAATACGTGAAAAA CCTGGGTTT	0.567	0.467	0.589	0.478	0.511	0.500	0.544	0.467
GDLNNTPGA (CNV1) GGAGACCTGAACAATACC CCAGGTGCA	0.778	0.767	0.633	0.722	0.711	0.756	0.633	0.767
VDIFSKPGV (CNV2) GGAGACCTGAACAATACC CCAGGTGCA	0.567	0.689	0.600	0.578	0.611	0.722	0.556	0.689

B

	GDVEENPGP (IMNV)	GDVESNPGP (IMNV)	HDIETNPGP (EMCV)	GDVELNPGP (PoRV-C)	GDIELNPGP (BoRV-C)	GDIESNPGP (BmCPV-1)	HDVEMNPGP (TMCV)	GDVESNPGP (FMDV)
EGGGKNPQW (LbTV) GAAGGCGGGGGAAAGAA CCCGCAGTGG	22.2%	22.2%	22.2%	22.2%	22.2%	22.2%	22.2%	22.2%
GDQSSNPQI (AToV) GGTGACCARTCTTCAAAC CCTCAGATC	44.4%	56.4%	33.3%	44.4%	44.4%	55.6%	33.3%	55.6%
IDSESNPSL (AATV) ATAGACTCGGAATCCAAC CCTTCATTG	44.4%	55.6%	44.4%	44.4%	44.4%	55.6%	44.4%	55.6%
QVIREKPGF (CYV) CAAGTAATACGTGAAAAA CCTGGGTTT	33.3%	22.2%	33.3%	22.2%	33.3%	33.3%	22.2%	22.2%
GDLNNTPGA (CNV1) GGAGACCTGAACAATACC CCAGGTGCA	44.4%	44.4%	33.3%	44.4%	44.4%	44.4%	33.3%	44.4%
VDIFSKPGV (CNV2) GGAGACCTGAACAATACC CCAGGTGCA	33.3%	44.4%	44.4%	33.3%	44.4%	55.6%	33.3%	44.4%

Fig. 4. Heatmaps presenting 2A-like ranking scores and sequence identity comparing pseudo 2A-sites with functional 2A-like sequences. (A) 2A-like ranking scores. (B) Similarity scores obtained from an alignment adopting BLOSUM62 matrix.

Previous studies concerning RNA viruses and 2 A-like peptides have reported that these sequences emerged independently during the evolution of viral families (Luke et al., 2008; Yang et al., 2017). Our results imply that this independence can be restricted at some level by the existence of precursor sequences herein called pseudo 2 A-sites. This would be in agreement with the phylogenetic proximity of the GLV-like and IMNV-like proposed groups, despite considerable differences in ORF1 structure. These differences have possibly started by events that altered the ancestral pseudo 2 A-site sequences in opposite directions, to promote or not the formation of functional 2 A-like sequences.

It was remarkable to note that only GLV, PMCV and PMCLV within the GLV-like group did not present pseudo 2 A-sites, confirming that somehow these sequences which would exist in their ancestors were lost. On the other hand, all IMNV-like members present functional 2 A-like sequences, a clear indication that this adaptation is an irreversible event. The subgroup formed by AToV, AATV, CYV, and CNV was in the "middle of the path", which preserved the ancestor sites which possibly originated the functional 2A-like sequences.

In agreement, our results also introduce that 2 A-like and pseudo 2 A-site sequences are submitted to positive and neutral selection, respectively, based in phylogenetic relationships and in the analysis of invariable sites (Table 2). Indeed, 2 A-like sequence cleavage capacity amplifies the viral adaptation potential, providing the advantage of releasing more than one protein from the same ORF. In this way, selective pressure can act on each protein individually, favoring adaptive changes in fewer generations and directly impacting viral infection mechanisms. Similarly, it has already been suggested that picornaviruses comprising multiple 2 A sequences have more complex infection mechanisms than others viruses of the same family (Yang et al., 2017). From this, it can be inferred that 2 A/2 A-like sequence presence in viral genomes can favor their fitness in complex multicellular systems.

4. Conclusion

The results presented herein reinforce considerable host and genome structural diversity within the *Totiviridae* family, especially considering the expanded *Giardiavirus* clade. Despite their considerable differences, especially in comparing ORF1 structure, the phylogenetic analysis revealed that GLV-like and IMNV-like groups present close evolutionary relationships, thus sharing the same common ancestor.

The structural diversity of ORF1 could be explained by the emergence of functional 2 A-like sequences from ancestral sequences described herein as pseudo 2 A-sites. These ancestral sequences reached the necessary configuration to codify functional 2 A-like sequences by random mutations, generating derived groups that could be observed in expanded *Giardiavirus* clade. In this way, some mutations in pseudo 2 A-sites may have prevented 2 A-like formation, originating the viruses GLV, PMCV, PMCLV and LbTV. Considering this scenario, AToV, AATV, CYV, CNV remained in the halfway point, conserving the ancestral sites with higher propensity to form functional 2 A-likes, while IMNV-like reaches the status of a different group from the emergence of functional 2 A-like sites. Thus, the pseudo 2 A-sites may be a key point in viral evolution, acting as a way to the increment or maintenance of genome complexity.

Funding information

This research received financial support from the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* - CNPq (grant no: 409378/2016-0)

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgements

The authors would like to thank the High Performance Computing Center (NPAD) of UFRN for provision of computational resources and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (*CAPES*) for financial support. Juliana G. S. de Lima and Diego G. Teixeira received scholarships from *CAPES*.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2018.10.011.

References

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.
Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., et al., 2014. BEAST 2: a software platform for bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.
Chen, S., Cao, L., Huang, Q., Qian, Y., Zhou, X., 2016. The complete genome sequence of a novel maize-associated totivirus. Arch. Virol. 161, 487–490.
Chen, V.B., Arendall 3rd, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., et al., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. 66, 12–21.
Colmant, A.M.G., Hobson-Peters, J., Bielefeldt-Ohmann, H., van den Hurk, A.F., Hall-Mendelin, S., Chow, W.K., Johansen, C.A., Fros, J., Simmonds, P., Watterson, D., Cazier, C., Etebari, K., Asgari, S., Schulz, B.L., Beebe, N., Vet, L.J., Piyasena, T.B.H., Nguyen, H.D., Barnard, R.T., Hall, R.A., 2017. A new clade of insect-specific flaviviruses from Australian *Anopheles Mosquitoes* displays species-specific host restriction. mSphere 2 (4) 17-e00262.
Dantas, M.D.A., Cavalcante, G.H.O., Oliveira, R.A.C., Lanza, D.C.F., 2016. New insights about ORF1 coding regions support the proposition of a new genus comprising arthropod viruses in the family Totiviridae. Virus Res. 211, 159–164.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27, 1164–1165.
Doherty, M., Sanganee, K., Kozlakidis, Z., Coutts, R.H.A., Brasier, C.M., et al., 2007. Molecular characterization of a totivirus and a partitivirus from the genus *Ophiostoma*. J. Phytopathol. (1986) 155, 188–192.
Donnelly, M.L.L., Hughes, L.E., Luke, G., Mendoza, H., ten Dam, E., et al., 2001a. The 'cleavage' activities of foot-and mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. J. Gen. Virol. 82, 1027–1041.
Donnelly, M.L.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., et al., 2001b. Analysis of aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J. Gen. Virol. 82, 1013–1025.
Drummond, A.J., Suchard, M., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res. 32, 1792–1797.
Fauver, J.R., Grubaugh, N.D., Krajacich, B.J., Weger-lucarelli, J., Lakin, S.M., et al., 2016. West African *Anopheles gambiae* mosquitoes harbor a taxonomically diverse virome including new insect-specific flaviviruses, mononegaviruses, and totiviruses. Virology 498, 288–299.
Galinier, R., Tetreau, G., Portet, A., Pinaud, S., Duval, D., et al., 2017. First characterization of viruses from freshwater snails of the genus Biomphalaria, the intermediate host of the parasite Schistosoma mansoni. Acta Trop. 167, 196–203.
Guo, L., Yang, X., Wu, W., Tan, G., Fang, S., et al., 2016. Identification and molecular characterization of Panax notoginseng virus A, which may represent an undescribed novel species of the genus Totivirus, family Totiviridae. Arch. Virol. 161, 731–734.
Hammond, T.M., Andrewski, M.D., Roossinck, M.J., Keller, N.P., 2008. Aspergillus mycoviruses are targets and suppressors of RNA silencing. Eukaryot. Cell 7, 350–357.
Haugland, Ø, Mikalsen, A.B., Nilsen, P., Lindmo, K., Thu, B.J., et al., 2011. Cardiomyopathy syndrome of Atlantic Salmon (*Salmo salar* L.) is caused by a double-Stranded RNA virus of the Totiviridae Family. J. Virol. 85, 5275–5286.
Isawa, H., Kuwata, R., Hoshino, K., Tsuda, Y., Sakai, K., et al., 2011. Identification and molecular characterization of a new nonsegmented double-stranded RNA virus isolated from Culex mosquitoes in Japan. Virus Res. 155, 147–155.
Jiang, Y., Zhang, T., Luo, C., Jiang, D., Li, G., et al., 2015. Prevalence and diversity of mycoviruses infecting the plant pathogen *Ustilaginoidea virens*. Virus Res. 195, 47–56.
Komatsu, K., Katayama, Y., Omatsu, T., Mizutani, T., 2016. Genome sequence of a novel victorivirus identified in the phytopathogenic fungus Alternaria arborescens. Arch. Virol. 161, 1701–1704.
Kondo, H., Hisano, S., Chiba, S., Maruyama, K., Andika, I.B., et al., 2016. Sequence and phylogenetic analyses of novel totivirus-like double-stranded RNAs from field-collected powdery mildew fungi. Virus Genes 213, 353–364.
Koyama, S., Sakai, C., Thomas, C.E., Nunoura, T., 2016. A new member of the family Totiviridae associated with arboreal ants (*Camponotus nipponicus*). Arch. Virol. 161, 2043–2045.
Koyama, S., Urayama, S., Ohmatsu, T., Sassa, Y., Sakai, C., et al., 2015. Identification, characterization and full-length sequence analysis of a novel dsRNA virus isolated from the arboreal ant *Camponotus yamaokai*. J. Gen. Virol. 96, 1930–1937.
Kozlakidis, Z., Covelli, L., Di Serio, F., Citir, A., Açikgöz, S., Hernández, C., Ragozzino, A., Flores, R., Coutts, R.H., 2006. Molecular characterization of the largest mycoviral-like double-stranded RNAs associated with Amasya cherry disease, a disease of presumed fungal aetiology. J. Gen. Virol. 87 (Pt 10), 3113–3117.
Lachnit, T., Thomas, T., Steinberg, P., 2016. Expanding our understanding of the seaweed holobiont: RNA viruses of the red alga Delisea pulchra. Front. Microbiol. 6, 1489.
Luke, G.A., de Felipe, P., Lukashev, A., Kallioinen, S.E., Bruno, E.A., et al., 2008. Occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes. J. Gen. Virol. 89, 1036–1042.
Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., et al., 2016. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 45 gkw1129.
Martinez, J., Lepetit, D., Ravallec, M., Fleury, F., Varaldi, J., 2016. Additional heritable virus in the parasitic wasp *Leptopilina boulardi*: prevalence, transmission and phenotypic effects. J. Gen. Virol. 97, 523–535.
Marzano, S.-Y., Nelson, B.D., Ajayi-Oyetunde, O., Bradley, C.A., Hughes, T.J., et al., 2016. Identification of diverse mycoviruses through metatranscriptomics characterization of the viromes of five major fungal plant pathogens. J. Virol. 90 JVI.00357-16.
Mor, S.K., Phelps, N.B.D., 2016a. Detection and molecular characterization of a novel piscine-myocarditis-like virus from baitfish in the USA. Arch. Virol. 161, 1925–1931.
Mor, S.K., Phelps, N.B.D., 2016b. Molecular detection of a novel totivirus from golden shiner (*Notemigonus crysoleucas*) baitfish in the USA. Arch. Virol. 161, 2227–2234.
Nerva, L., Ciuffo, M., Vallino, M., Margaria, P., Varese, G.C., et al., 2016. Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. Virus Res. (219), 22–28.
Nibert, M.L., 2007. '2A-like' and 'shifty heptamer' motifs in penaeid shrimp infectious myonecrosis virus, a monosegmented double-stranded RNA virus. J. Gen. Virol. 88, 1315–1318.
Niu, Y., Zhang, T., Zhu, Y., Yuan, Y., Wang, S., et al., 2016. Isolation and characterization of a novel mycovirus from *Penicillium digitatum*. Virology 494, 15–22.
Notredame, C., Higgins, D.G., Heringa, J., 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.
Ong, J.W.L., Li, H., Sivasithamparam, K., Dixon, K.W., Jones, M.G.K., et al., 2018. Novel and divergent viruses associated with Australian orchid-fungus symbioses. Virus Res. 244, 276–283.
Osaki, H., Sasaki, A., Nomiyama, K., Tomioka, K., 2016. Multiple virus infection in a single strain of *Fusarium poae* shown by deep sequencing. Virus Genes 52, 835–847.
Poulos, B.T., Tang, K.F.J., Pantoja, C.R., Bonami, J.R., Lightner, D.V., 2006. Purification and characterization of infectious myonecrosis virus of penaeid shrimp. J. Gen. Virol. 87, 987–996.
Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E.H., Man, O., Beckmann, J.S., Silman, I., Sussman, J.L., 2005. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics 21 (16), 3435–3438 Epub 2005 Jun 14.
Quito-Avila, D.F., Alvarez, R.A., Mendoza, A.A., 2016. Occurrence of maize lethal necrosis in Ecuador: a disease without boundaries? Eur. J. Plant Pathol. 146, 705–710.
Rodriguez-Cousiño, N., Gomez, P., Esteban, R., 2017. Variation and distribution of L-A helper totiviruses in *Saccharomyces sensu stricto* yeasts producing different killer toxins. Toxins 9, e313.
Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738.
Rowley, P.A., Ho, B., Bushong, S., Johnson, A., Sawyer, S.L., 2016. XRN1 is a species-specific virus restriction ractor in yeasts. PLoS Pathog. 12, e1005890.
Tang, L., Hu, Y., Liu, L., Wu, S., Xie, J., 2015. Genomic organization of a novel victorivirus from the rice blast fungus Magnaporthe oryzae. Arch. Virol. (160), 2907–2910.
Tuomivirta, T.T., Hantula, J., 2005. Three unrelated viruses occur in a single isolate of *Gremmeniella abietina* var. abietina type A. Virus Res. (110), 31–39.
Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., Barton, G.J., 2009. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191.
Wickner, R.B., Ghabrial, S.A., Nibert, M.L., Patterson, J.L., Wang, C.C., 2012. Family - totiviridae. Part II: the viruses. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. Academic Press, Waltham, MA, pp. 639–650.
Wu, B., Zhang, X., Gong, P., Li, M., Ding, H., 2016a. Eimeria tenella: a novel dsRNA virus in *E. Tenella* and its complete genome sequence analysis. Virus Genes 52, 244–252.
Wu, M., Deng, Y., Zhou, Z., He, G., Chen, W., et al., 2016b. Characterization of three mycoviruses co-infecting the plant pathogenic fungus *Sclerotinia nivalis*. Virus Res. 223, 28–38.
Wu, Q., Luo, Y., Lu, R., Lau, N., Lai, E.C., et al., 2010. Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. Proc. Natl. Acad. Sci. U. S. A. 107, 1606–1611.
Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. J. Hered. 92, 371–373.
Xin, C., Wu, B., Li, J., Gong, P., Yang, J., et al., 2016. Complete genome sequence and evolution analysis of Eimeria stiedai RNA virus 1, a novel member of the family Totiviridae. Arch. Virol. 161, 3571–3576.
Yamada, K.D., Tomii, K., Katoh, K., 2016. Application of the MAFFT sequence alignment program to large data - reexamination of the usefulness of chained guide trees. Bioinformatics 32, 3246–3251.
Yang, X., Cheng, A., Wang, M., Jia, R., Sun, K., et al., 2017. Structures and corresponding functions of five types of picornaviral 2A proteins. Front. Microbiol. 8, 1373.
Yang, X., Zhang, Y., Ge, X., Yuan, J., Shi, Z., 2012. A novel totivirus-like virus isolated from bat guano. Arch. Virol. 157, 1093–1099.

Zangger, H., Hailu, A., Desponds, C., Lye, L.-F., Akopyants, N.S., et al., 2014. *Leishmania aethiopica* field isolates bearing an endosymbiontic dsRNA virus induce pro-inflammatory cytokine response. PLoS Negl. Trop. Dis. 8, e2836.

Zhai, Y., Attoui, H., Mohd Jaafar, F., Wang, H.Q., Cao, Y.X., et al., 2010. Isolation and full-length sequence analysis of *Armigeres subalbatus totivirus*, the first totivirus isolate from mosquitoes representing a proposed novel genus (Artivirus) of the family Totiviridae. J. Gen. Virol. 91, 2836–2845.

Zhai, L., Hong, N., Zhang, M., Wang, G., 2015. Complete dsRNA sequence of a novel victorivirus isolated from the pear stem wart fungus *Botryosphaeria dothidea*. Arch. Virol. 160, 613–616.

Zhang, R.J., Jie, Z., Shang, H.H., Pan, T., Zhu, J.H., et al., 2015. The complete nucleotide sequence and genomic organization of a novel victorivirus with two non-overlapping ORFs, identified in the plant-pathogenic fungus *Phomopsis vexans*. Arch. Virol. 160, 1805–1809.

Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9, 40.

Zheng, L., Lu, X., Liang, X., Jiang, S., Zhao, J., et al., 2017. Molecular characterization of novel totivirus-like double-stranded RNAs from *Puccinia striiformis* f. Sp. *tritici*, the causal agent of wheat stripe rust. Front. Microbiol. 8, 1960.

Zhong, J., Cheng, C.Y., Gao, B.D., Zhou, Q., Zhu, H.J., 2017. Mycoviruses in the plant pathogen *Ustilaginoidea virens* are not correlated with the genetic backgrounds of its hosts. Int. J. Mol. Sci. 18, 963.

Zhong, J., Hua, X., Jun, L., Zhu, Z., 2014a. Detection and sequence analysis of two novel co-infecting double-strand RNA mycoviruses in Ustilaginoidea virens. Arch. Virol. 159, 3063–3070.

Zhong, J., Zhou, Q., Hu, Y., Jian, H., Da Gao, B., 2016. Molecular identification of a novel victorivirus from the phytopathogenic fungus Nigrospora oryzae. Virus Genes (52), 156–159.

Zhong, J., Zhou, Q., Lei, X.H., Chen, D., Shang, H.H., et al., 2014b. The nucleotide sequence and genome organization of two victoriviruses from the rice false smut fungus *Ustilaginoidea virens*. Virus Genes 48, 570–573.