Virus Research 313 (2022) 198727

Contents lists available at ScienceDirect

Virus Research

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

Readdressing the genetic diversity and taxonomy of the *Mesoniviridae* family, as well as its relationships with other nidoviruses and putative mesonivirus-like viral sequences

Paulo Morais [a,*], Nídia S. Trovão [b], Ana B. Abecasis [c], Ricardo Parreira [a]

[a] Unidade de Microbiologia Medica, Instituto de Higiene e Medicina Tropical (IHMT), Universidade Nova de Lisboa (NOVA), Lisboa, Portugal/Global Health and Tropical Medicine (GHTM), Lisboa, Portugal

[b] Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

[c] Unidade de Saúde Pública Internacional e Bioestatística, Instituto de Higiene e Medicina Tropical (IHMT), Universidade Nova de Lisboa (NOVA), Lisboa, Portugal/Global Health and Tropical Medicine (GHTM), Lisboa, Portugal

ARTICLE INFO

Keywords:
Mesonivirus
Nidovirales
Taxonomy
Genetic diversity
Phylogenetic analysis

ABSTRACT

Research on the recently established *Mesoniviridae* family (Order *Nidovirales*), RNA genome insect-specific viruses, has been steadily growing in the last decade. However, after the last detailed phylogenetic characterization of mesoniviruses in 2014, numerous new sequences, even in organisms other than mosquitos, have been identified and characterized.

In this study, we analyzed nucleotide and protein sequences of mesoniviruses with a wide range of molecular tools including genetic distance, Shannon entropy, selective pressure analysis, polymorphism identification, principal coordinate analysis, likelihood mapping and phylodynamic reconstruction. We also sought to revaluate new mesoniviruses sequence positions within the family, proposing a taxonomic revision.

The different sub-lineages of mosquito mesoniviruses sequences presented low sequence diversity and entropy, with incongruences to the existing taxonomy being found after an extensive phylogenetic characterization. High sequence discrepancy and differences in genome organization were found between mosquito mesoniviruses and other mesoniviruses, so their future classification, as other meso-like viruses that are found in other organisms, should be approached with caution.

No evidence of frequent recombination was found, and mesonivirus genomes seem to evolve under strong purifying selection. Insufficient data by root-to-tip analysis did not yet allow for an adequate phylogeographic reconstruction.

1. Introduction

The Order *Nidovirales* comprises a genetically diverse assemblage of enveloped, approximately spherical viruses with linear single-stranded, positive-sense, and polyadenylated RNA genomes, that can infect a wide range of hosts, from mammals to insects. According to the International Committee on Taxonomy of Viruses (ICTV), they are taxonomically (mid-2021) distributed in eight suborders and 14 families (https://talk.ictvonline.org/taxonomy/), including the well-studied *Arteriviridae, Coronaviridae*, and *Roniviridae,* as well as the more recently established *Mesoniviridae* family (Vasilakis *et al.*, 2014).

Within the Order *Nidovirales*, mesoniviruses were the first known to infect insects. Their detailed description was initiated in 2011 with the characterization of the Cavally (CAVV) and Nam Dinh (NDiV) viruses, isolated from *Culex* mosquitos, collected in Cote d'Ivoire and Vietnam, respectively (Zirkel *et al.*, 2011; Nga *et al.*, 2011). Since then, mesoniviruses have been isolated from mosquitos collected in the Americas (Charles *et al.*, 2018), Asia (Wang *et al.*, 2017), Africa (Diagne *et al.*, 2020), and Oceania (Warrilow *et al.*, 2014), suggesting a global distribution. Like insect-specific flaviviruses (Blitvich and Firth, 2015) and mosquito-associated bunyaviruses (Marklewitz *et al.*, 2013), mesoniviruses are considered some of the most predominant RNA genome insect-specific viruses (ISVs) (Vasilakis et al., 2014). While they have repeatedly been isolated from naturally infected mosquitoes, they do not appear to infect vertebrates (Blitvich and Firth, 2015). Nonetheless, their isolation from *Aphis citricidus* aphids collected in 2012 in China

* Corresponding author.
E-mail address: a21000757@ihmt.unl.pt (P. Morais).

https://doi.org/10.1016/j.virusres.2022.198727
Received 25 November 2021; Received in revised form 2 March 2022; Accepted 3 March 2022
Available online 5 March 2022
0168-1702/© 2022 Elsevier B.V. All rights reserved.

suggests that the host range of mesoniviruses might go beyond that which is currently known (Chang *et al.*, 2020). Furthermore, a meso-like virus has already been detected in Italy in *Leveillula taurica*, an obligate fungal pathogen (accession number MN609866).

The genomes of mesoniviruses of approximately 20 kb are organized into multiple open-reading frames (ORFs). The most frequently found organization is ORF1a-ORF1b-ORF2a-ORF2b-ORF3a-ORF3b-ORF4, but exceptions do exist (e.g. the Meno virus does not encode ORF4). The 5′ region of the genome encodes two polyproteins (ORF1a and ORF1b), the expression of which is controlled by ribosomal frameshift followed by proteolytic processing (Vasilakis et al., 2014), and their products are suggested to be involved in the regulation of gene expression, polyprotein processing, and genome replication and transcription. The 3′ region of the genome includes smaller ORFs that encode structural proteins. Apart from ORF1a and ORF1b, the number of small ORFs varies among different viruses in the Order *Nidovirales* (Gorbalenya *et al.*, 2006).

The latest update from the ICTV regarding the *Mesoniviridae* family (March 2021, available at https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/308/mesoniviridae), acknowledges 1 single genus (*Alphamesonivirus*) and 8 subgenera. *Namcalivirus* is represented by both the Alphamesonivirus 1 species (comprising most mesoniviruses isolated to date), and the Alphamesonivirus 10 species (which includes the Dianke virus). Other genera encompass only one other viral type. For example, the Ofaie virus (OFAV) and the Casuarina virus (CASV) are currently the sole representatives of the *Ofalivirus* (Alphamesonivirus 6) and *Casualivirus* (Alphamesonivirus 4) genera, respectively. In addition, several recently discovered mesoniviruses [e.g. the Odorna virus (OdoV)], remain unclassified.

Considering the recent pandemic spread of SARS-CoV-2 coronavirus the interest in the study of mesoniviruses has increased (Lai *et al.*, 2020). While distantly related to coronaviruses and mostly restricted to mosquitoes, their study might hold crucial information regarding the evolution of the viruses within the Order *Nidovirales*, as they may have evolved in arthropods (Nga *et al.*, 2011). However, while the genomic and phylogenetic characterization of mesonivirus has lastly been addressed in a comparative dating from 2014 (Vasilakis *et al.*, 2014), since then, the isolation of multiple mesoniviruses prompted us to revaluate their position within the family. Furthermore, the recent discovery of a meso-like virus in organisms other than mosquitos might hold new information regarding their phylogenetic relationship with other mesoniviruses. In this report, we will also discuss the conditions required for a potential future phylogeographic analyses of this *taxa*.

2. Materials and methods

2.1. Dataset preparation and sequence alignments

The compilation of the different nucleotide (nt) and amino acid (aa) sequence datasets used in this work was based on the selection of complete genome sequences available at GenBank in 01/07/2021. These were either directly identified via their respective accession numbers, or indirectly singled out as a product of similarity searches using BLASTn.

All sequences corresponding to complete genomes available were downloaded, and additional information including GenBank accession number, host species, geographic origin, and collection date was obtained. When available, information on genomic coding-capacity (ORF organization) and their respective sequences were also collected. Furthermore, for comparative purposes, representative datasets containing ORF1ab nt and aa sequences of viruses from other families in the Order *Nidovirales* (corresponding to the most conservative coding region between them) were also constructed.

Multiple alignments of complete nt and aa sequences were performed using the iterative G-INS-I method as implemented in MAFFT v. 7 (Katoh and Standley, 2013), followed by their edition using GBlocks (Castresana, 2000), allowing for less strict flanking positions in the obtained multiple sequence alignments (MSA). These were systematically verified to ensure the correct alignment of homologous codons using BioEdit 7.0.5 (Hall, 1999). Additional alignments were also constructed for different ORFs identified in the *Mesoniviridae* family that included ORF1a, ORF1b, ORF2a/spike, ORF3a, and ORF4, as well as the viral RNA-dependent RNA polymerase (RdRp). Multiple alignments of ORF1ab aa sequences from different families in the *Nidovirales* Order were performed similarly using MAFFT iterative L-INS-I option, followed by a new alignment using the G-INS-I method.

2.2. Assessment of the temporal and phylogenetic signals of different mesonivirus sequence datasets

The evolutionary information contained in all used sequence datasets (phylogenetic signal) was assessed by Likelihood Mapping (Strimmer and von Haeseler, 1997) using TREE-PUZZLE v5.3 (Schmidt et al., 2002). Datasets with totally resolved quartets values of over 90% were considered of high phylogenetic resolving power.

A visual inspection of the degree of temporal signal (i.e. signal for divergence accumulation over the sampling time interval) in the complete genome nt datasets (as well as for the RdRp and S protein-coding sequences) for all mesoniviruses was carried out using an exploratory linear regression approach assuming the topology of a Maximum Likelihood (ML) tree, estimated under a non-clock (unconstrained) and the GTR+Γ+I substitution model using IQ-TREE (Trifinopoulos et al., 2016). Root-to-tip divergences were plotted as a function of sampling time using the TempEst v. 1.5.3 (Rambaut et al., 2016).

2.3. Phylogenetic analyses using maximum likelihood

Phylogenetic reconstructions of full-length genomic nt and ORF-specific nt datasets and specific aa sequences (RdRp and S datasets) were performed based on the maximum likelihood optimization criterion, using the GTR+Γ+I model and Whelan And Goldman (WAG) model, respectively, as suggested by IQ-TREE (Trifinopoulos et al., 2016), which was also used for tree building. The stability of the obtained tree topologies was assessed by bootstrapping and using the aLRT (approximate likelihood ratio test) with 1000 re-samplings of the original sequence data.

2.4. Genetic diversity and protein primary sequence analyses

The estimation of genetic distance values (corrected with the Kimura-2P formula) was carried out using MEGAX (Kumar et al., 2018). Heat maps were calculated based on pairwise evolutionary distances obtained using the Heatmapper webserver (Babicki *et al.*, 2016), while box plots were drawn with Microsoft Excel. Visualization of viral genomic organization was performed using Open Reading Frame Finder (available in https://www.ncbi.nlm.nih.gov/orffinder/), while the SMART webserver (Letunic and Bork, 2018) was used for the identification and annotation of genetically mobile domains. The presence of conserved domains in viral protein sequences was investigated using CD-Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Remote homology detection and structure prediction was analyzed using HHblits and Hpred, as well as sensitive sequence searching by HHMER (Zimmermann *et al.*, 2018). Several bioinformatic tools were employed to investigate ORFs with unknown function, including computation of molecular weight and theoretical isoelectric point (pI) via ProtParam (https://web.expasy.org/protparam/), analysis of hydropathicity by ProtScale (https://web.expasy.org/cgi-bin/protscale/protscale.pl), prediction of transmembrane helices via TMHMM v2.0 (http://www.cbs.dtu.dk/services/TMHMM), prediction of n-glycosylation sites (http://www.cbs.dtu.dk/services/NetNGlyc) and o-glycosylation sites (http://www.cbs.dtu.dk/services/NetOGlyc),

signal sequence search by SignalP (http://www.cbs.dtu.dk/services/SignalP) and protein subcellular localization prediction by DeepLoc (http://www.cbs.dtu.dk/services/DeepLoc). The detection of repeats in protein sequences was carried out with RADAR (https://www.ebi.ac.uk/Tools/pfa/radar/).

The analyses of selective pressure on individual sites of codon alignments were carried out using the Single Likelihood Ancestor Counting (SLAC), the Fixed Effects Likelihood methods as implemented in Datamonkey (Kosakovsky Pond and Frost, 2005), or the SNAP tool (http://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html) that explores a simple method for calculation of synonymous and non-synonymous substitutions (Nei and Gojoborit, 1986). The degree of variability at each amino acid position in multiple alignments of single ORF aa sequences was evaluated based on the Shannon entropy function using Entropy (Shannon entropy-one option, available at http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html). Finally, principal coordinate analysis was carried out using PCOORD (http://www.hiv.lanl.gov/content/sequence/PCOORD/PCOORD.html). Possible recombination events were investigated using the Recombination Detection Program 4 (RDP4) software (Martin et al., 2015).

2.5. Comparative analysis with virus from other Nidovirales families

For comparative analyses of mesonivirus genomic sequences with those of other nidoviruses, overall mean distances, assessment of phylogenetic signals, and selective pressure analyses were performed for families in the *Nidovirales* Order with higher representation in the genomic sequence databases (*Arteriviridae, Coronaviridae,* and *Tobaniviridae*), focusing on the most conserved coding region among them (ORF1ab).

To assess the relationship between different families in the *Nidovirales* Order, phylogenetic reconstructions were carried out using multiple sequence alignments of RdRp aa sequences as described in Section 2.3.

3. Results

3.1. Comparative genome organization analyses

A total of 47 full-length mesonivirus genomic sequences, downloaded from the public genomic databases, were aligned, and analyzed. These included both those that, until 2020, had been only identified in multiple species of mosquitoes ($n = 44$), being frequently associated with *Culex sp.* or *Aedes sp.* In addition, this dataset also included three meso-like viral sequences that had recently been identified in hosts other than mosquitoes. These comprised those of meso-like viruses isolated from *Aphis citricidus* aphids (*Aphis citricidus* meson-like virus, AcMSV), from *Thrips tabaci* thrips (Insect metagenomics mesonivirus 1, Immeso1; Chiapello et al., 2021), as well as from a fungal pathogen, *Leveillula Taurica* (*Leveillula taurica* associated alphamesonivirus 1, temporarily abbreviated as LtM). All these have been listed in Supplementary Table 1. Additionally, for phylogenetic and other comparative analyses, ORF1ab aa sequences were also compiled for viruses in other families in the *Nidovirales* Order, and these have been included in Supplementary Table 2. Alongside the full-length genome datasets, other datasets including the nt and aa sequences of all mesonivirus identifiable ORFs (of the sequences listed in Supplementary Table 1) were also constructed.

Also, as suggested by Gorbalenya et al. in 2006, and as corroborated here in Supplementary Fig. 1, the number of ORFs identified in viral genomes from viruses allocated to different families in the *Nidovirales* Order is substantially different. Viruses from the *Mesoniviridae* family display smaller genomes with 4 to 7 ORFs, as similarly observed in the *Tobaniviridae* and *Medoniviridae* families. In contrast, viruses from the *Arteriviridae* and *Coronaviridae* families have a larger number of ORFs, up to 12.

A comparative analysis of the organization predicted for the different mosquito mesoniviruses (MM) genomes (Supplementary Fig. 1) indicated overall conserved synteny, with only those of meno-, kadiweu- and ofaieviruses missing an identifiable ORF4. As no complete genomic sequence have yet been made available for OdoV, a prediction of its genomic structures remains incomplete. A comparison between the genome organization of MM and other mesoniviruses (OM) revealed substantial differences, especially considering their similar genome sizes (excluding the 3'-poly [A] tail, they range from 19,209 nt for Immeso1 to 20,626 nt for AcMSV). All these viruses are suggested to use ribosomal frameshifting for translational control of the expression of non-structural proteins, as revealed by the consistent overlap between ORF1a and ORF1b, while ORF2a (surface spike) encodes the S glycoprotein. As expected, and considering that the mature products of ORF1a and 1b are usually involved in the control of essential steps of the viral replication cycle that include genomic replication, transcription, RNA capping and polyprotein processing, the genomic organization appears similar when MM and OM are compared (although smaller in size in LtM), including most conserved domains and other so-called genetically mobile domains (i.e. transposable elements; Vasilakis et al., 2014) (Fig. 1). Only Immeso1 does not seem to possess a coiled coil motif in ORF1a, while displaying a zinc finger domain which, however, is not shared by other mesoniviruses. Unfortunately, the available LtM genomic sequence appears to be incomplete, with only the full sequence of ORF1a and a partial sequence of ORF1b currently available. While the ORFs at the 3' half of the genome of MM were similar (except for ORF4), the number and organization of the ORFs identified in that same region of the genomes of OM are different. These ORFs (identified as ORFs x1-3 and ORFs y1-3 in Fig. 1) seem to encode putative products that, in most cases, share no easily identifiable homology with other viral proteins, nor do they display readily recognizable conserved domains (as defined by CD-BLAST analysis) associated with a particular biological function or protein family. Sequence searches regarding both y1 and y3 did not return any putative matches with homologues in the sequence databases, not even when remote homology detection methods were used (HHblits, HHpred, or HMMER). However, y1 is predicted as a 22kDa, basic (pI 9.9) and hydrophilic, while y3 is also small (12kDa) and basic but mostly hydrophobic. Furthermore, while multiple O-glycosylation sites were predicted in y1, none have been predicted for y3. On the other hand, y2 is larger (98 kDa) and mostly neutral. However, remote homology detection tools indicated a 96% probability match between the highly basic (pI=10.5) product of ORFx2 (between amino acids 108 and 216) and the putative nucleocapsid protein of the Kadiweu virus (Alphamesonivirus 7), while part the product of ORFy2 aligned with the ORF2a protein encoded by NDiV (positions 391 and 906). In addition, the larger ORF found in the genomes of the *Aphis citricidus* meson-like virus and the *Thrips tabaci* associated mesonivirus (ORFs x3 and y2 in Fig. 1) encode putative proteins with 3 (y2) or 6 (x3) transmembrane helixes and multiple targets for N-glycosylation, which are features frequently found in integral viral envelope proteins. Finally, the putative ORF4, which is encoded by the genomes of mosquito mesoniviruses, is highly conserved (96.3% identity conservation among MM). It encodes a small (approximately 5kDa), basic (pI=9.6) hydrophilic protein, with no glycosylated amino acids, transmembrane helices, or signal peptide sequences (the latter found in the product of *Aphis citricidus* meson-like virus ORFx1), or conserved domains. Up to the present, its function remains unknown.

In addition, insertions blocks in ORF1a have been described in a handful of mesonivirus (Kamphaeng Phet, KPhV; Karang Sari virus, KSaV; Bontang Baru, BBaV; Vasilakis et al., 2014), but among the more recently identified MM and OM, the Dak Nong virus (DkNV) also revealed those same type of insertions. Two of these insertion blocks are larger (approximately 570 nt) than the other (approximately 170 nt) but despite their size difference they partially align at their 5′ ends. All these insertions extend the coding capacity of ORF1a, with the larger of these two insertions, found in the genomes of the BBaV and KSaV viruses, apparently coding for hydrophilic peptides of approximately 190 aa

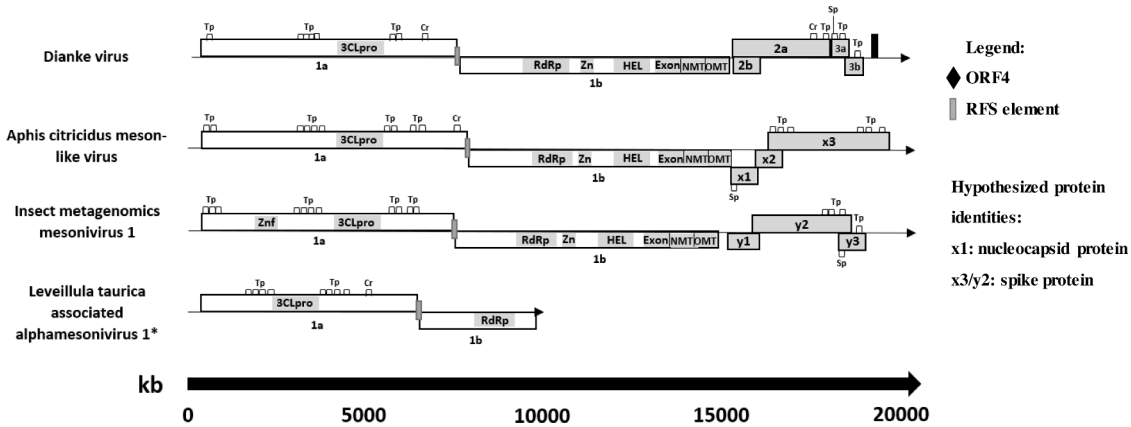

Fig. 1. Schematic representation of the genomic organization of mosquito mesoniviruses (Dianke virus, accession number MN622133, used as an example) and other mesoniviruses: Aphis citricidus meson-like virus (AcMSV, accession number MN961271); Insect metagenomics mesonivirus 1 (Immeso1, accession number MN714662); Leveillula taurica associated alphamesonivirus 1 (LtM, accession number MN609866); Znf = Zinc finger; 3CLpro = 3C-like protease; RdRp = RNA-dependent RNA polymerase; Zn = Zinc-binding domain; HEL = Helicase; Exon = 3′-5′ exoribonuclease; NMT = N7-methyltransferase; OMT = Cap-0 specific (nucleoside-2′-O-)-methyltransferase; Tp = Transmembrane region; Cr = Coiled coil region; Sp = Signal peptide. * - Leveillula taurica associated alphamesonivirus 1 whole genome sequence still not available. ORFs at the 3' half of the genome for other mesoniviruses identified as x1-x3 (for AcMSV) and y1-y3 (for Immeso1), with most recognizable putative proteins for each (when available) displayed.

characterized by two types of partially repeated sequences. One of these is repeated 3 to 4 times at the N-terminal section, while the other (in two copies), can be identified at the peptide's C-terminus. Homology searches did not reveal sequence similarities or amino acid motifs that might indicate their putative function, but the sequence encoded by the BBaV is characterized by the presence of a PKR13108 superfamily sequence motif found in prolipoprotein diacylglyceryl transferases, found in bacteria of the *Corynebacterineae* family (E=7.64e^{-04}).

3.2. Genetic diversity analyses

Overall mean genetic distances for MM were calculated for both the complete genome as well as for each ORF-specific genomic region (Supplementary Table 3). Furthermore, both genetic distances between all MM full-length genomes as well as for two ORF-specific genomic regions (RdRp and S) were also calculated (Supplementary Table 4). The overall mean distance (complete genome) between all sequences was 0.15. The region encompassing the ORF4 gene was the viral genomic region with the lowest mean genetic distance value (0.04), while the ORF1a region held the highest (0.17). Using ORF1ab aa sequences as a reference, analyses of datasets including up to three different sequences genus/subgenus selected from the viral families with a larger representation in the public databases (*Mesoniviridae, Coronaviridae, Arteriviridae* e *Tobaniviridae*), overall mean genetic distance values were 0.07, 0.34, 0.43 and 0.50, respectively. Pairwise evolutionary distances (PEDs) were calculated between all RdRp aa mesonivirus sequences (Supplementary Table 5), with heat maps designed to visualize inter-sequence genetic diversity, and box-and-whisker graphs used to visualize all distances between mesoniviruses from the same species, between mesonivirus from different species and also between MM and OM. These analyses clearly highlight the difference between MM and OM, as seen in Supplementary Fig. 2. Substantial differences in PEDs between all groups analyzed were also highlighted using box-and-whisker graphs (Supplementary Fig. 3). While, and as expected, higher PEDs were obtained when protein sequence comparisons included viral sequences from different species (as opposed to intra virus species comparisons), with values always below the 96.8% protein sequence identity used as cut-off (Vasilakis *et al.*, 2014), but this distance was far considerably pronounced when comparing MM to OM, with identity values under 70%. Two groups of mesoniviruses sequences, BBaV and KPhV, clearly exceed the cut-off value, with identity values below 80%, which suggests they should correspond to new species.

Shannon entropy is a quantitative measure of uncertainty in a dataset of nucleotide or amino acid sequences, and it may be considered as a measure of variation in DNA and protein sequence alignments for assessment of genetic diversity in a cross-sectional sense. When applied to the analysis of MM mesonivurus sequences, Shannon entropy calculations showed low values for all mesonivirus ORF-coding sequences. However, statistically higher entropy values were calculated for ORF1a, especially when compared to other genomic regions, while ORF3 showed the lowest entropy. Other families in the *Nidovirales* Order had consistently higher entropy values when compared with the *Mesoniviridae* family (Supplementary Fig. 4).

3.3. Phylogenetic analyses of mosquito mesoniviruses

For different datasets of nt mesonivirus sequences, likelihood mapping analyses were performed to calculate their respective phylogenetic signals (Table 1). The obtained results showed an overall high percentage (>90%) of the totally resolved sequence quartets (assessing the topologies of 10,000 quartet replicates) obtained for the complete genome, ORF1a, ORF1b and ORF2a, as well as the specific RdRp-coding sequence, while lower values were obtained for ORF3a (81.8%) and ORF4 (74.1%). These results indicate that most phylogenetic reconstructions based on the analysis of alignments for any viral ORFs other than ORF3 and 4 could be used to produce unambiguous trees. Overall, phylogenetic reconstructions using full-length genome sequences from viruses allocated to different families in the *Nidovirales* Order consistently presented high phylogenetic signals (Supplementary Table 6). However, while the great majority of the constructed datasets revealed consistent high phylogenetic signals, standard linear regression exploration of root-to-tip distances as a function of sampling time (Supplementary Table 7) carried out to establish to what extent the *Mesoniviridae* family contained detectable signal for sequence divergence throughout the sampling time intervals, showed negative slope and correlation coefficient values, even after an extensive root-to-tip analysis and the removal of outlier sequences. This observation extended for both the full-genome, RdRp and S protein-coding sequence comparisons, as well as when analyzing only the Alphamesonivirus 1 species or all MM and OM at once (indicated as all mesoniviruses in Supplementary Table 7). As such, at the present, only the investigation of phylogenetic relationships using ML trees is possible, while potential temporal and phylogeographical analyses using a Bayesian phylodynamic framework await the description of future new mesonivirus

Table 1
Phylogenetic signal of mesonivirus sequence datasets.

Datasets	Full-legnth genome	ORF1a	ORF1b	RdRp	ORF2a (S)	ORF3a	ORF4
Totally resolved quartets	99.3%	98.4%	98.7%	94,4%	96.8%	81.8%	74.1%
Partially resolved quartets	0.6%	0.8%	0.8%	2.7%	1.2%	3.7%	1.7%
Unresolved quartets	0.1%	0.9%	0.4%	2.9%	2.1%	14.4%	24.2%

sequences.

Selective pressure analyses were carried by estimating omega (ω, i.e. dN/dS or the frequency of non-conservative-to-conservative substitutions ratio) values using concatenated ORF1a/ORF1b/ORF2/ORF3/ORF4 coding-sequence datasets, as well as for each one of the individual ORFs using three different methods (SLAC, FEL and SNAP). No significant differences were found between all ORFs, apart from ORF4 (which seems to be the only region under diversifying selection, with ω values over 1, even though it is the genomic region with lowest genetic diversity), with all ω values being very low (Supplementary Table 8a), with site-specific selection analysis revealing high percentage of negatively selected sites, as well as very low percentage of positively selected sites. Comparative analyses with other families in the *Nidovirales* Order were performed using their most conservative region among their genomes, the ORF1ab region (Supplementary Table 8b). Unlike mesonivirus sequences, those from coronaviruses, tobaniviruses and arteriviruses displayed higher ω values, always higher than 1, with lower percentage of negatively selected sites and higher percentage of positively selected sites.

Recombination events are common among viruses classified into the *Nidovirales* Order (Gorbalenya *et al.*, 2006) and has been shown to affect the evolution of some of its best studied members (Hon *et al.*, 2008; Boni *et al.*, 2020). Since no previous assessment of whether these events affect the evolution of mesoniviruses had ever been performed, we investigated whether this would extend to mesoniviruses using the RDP4 software. A full genome scan (using all detection methods implemented in RDP4) was performed, and while many minor recombination events were detected, only one potential recombination event was strongly supported by the RDP4 software, regardless of the recombination detection method used. This event seems to have been involved in the genesis of NDiV (accession number KF771866), as its genome appears to have resulted from the recombination of two distinct mesoniviruses, with NgeV (accession number MF176279) and OdoV (LC497422), or viruses very similar to them, suggested as the parental sequences. Due to the apparent mosaic nature of the NDiV sequence, it was removed from further phylogenetic analyses.

Phylogenetic reconstructions were based on the analyses of mesoniviruses genomic regions with high phylogenetic signals (Table 1). We focused our analyses on non-mosaic full-length genome sequences (dataset with the highest phylogenetic signal), as well as two others comprising different ORF-specific datasets with higher signal (ORF1b and ORF2a/S) which encode very different types of viral proteins. However, instead of analyzing the whole of the ORF1b-coding sequence, we sought to focus our analyses exclusively on the RdRp coding sequences, which not only displays high phylogenetic signal, but especially because it is, by far, the mesonivirus genomic sequence most frequently found in public databases.

Phylogenetic reconstructions carried out using either the ORF2a/S-coding region or the full-length genomic sequence translated into similar results (Supplementary Fig. 5 for ORF2a/S and Supplementary Fig. 6 for full-length). When the current taxonomy of mesoniviruses (according to the latest update on ICTV) is superimposed to the topology of these trees and to the mesonivirus species demarcation criteria defined by Vasilakis *et al.* in 2014 (96.8% protein sequence identity), evident discrepancies where found when the topologies of the complete genome/S and RdRp trees were compared (compare Supplementary Fig. 5 vs. Fig. 2). Even as the trees appeared to be topologically similar, they were not identical. For example, CAVV was placed in the lineage defining the Alphamesonivirus 1 species only in the RdRp tree, and this association seemed to be supported by all PCOORD analyses. In addition, the monophyletic group that included DkNV and KPhV sequences is indicated, in the RdRp tree (and is supported by PCOORD), as sharing direct ancestry and forming a robust monophyletic clade with the lineage that clusters KSaV and BBaV, when this is not seen in the S-protein tree. These results indicate that while the mesonivirus genome or the RdRp and S regions may be used for phylogenetic analysis, some topological discrepancies are seen depending on the region used. Clearly, if tree topologies are considered to aid taxonomic decisions, even slightly different topologies may impact viral taxonomy.

Since species demarcation criteria for nidoviruses have been most focused on highly conserved RdRp aa sequences (Cowley and Walker, 2014), and since mesonivirus species demarcation have previously been focused on the analysis of concatenated regions of highly conserved domains within ORF1ab (Vasilakis *et al.*, 2014), to evaluate how the species demarcation criteria would affect mesoniviruses classification, we focused our attention on the RdRp aa tree (Fig. 2). When phylogenetic relationships among MM were superimposed to the nomenclature scheme current defining the mesonivirus taxonomy, some clashes between tree topology and previous taxonomy assignments become apparent. KPhV and BBaV are both assigned as members of the Alphamesonivirus 1/AMV1 species and of the *Namcalivirus* subgenus in NCBI's taxonomy browser (Schoch et al., 2020), which would mean, by the obtained tree topology, that namcaliviruses are paraphyletic. However, previous studies never did place KPhV and BBaV into that specific Alphamesonivirus species (Vasilakis *et al.*, 2014; Wang *et al.*, 2017; Newton *et al.*, 2020), with Vasilakis *et al.* (2014) even suggesting, by PED analysis, that these two species should be considered as separate species. Indeed, our analysis did confirm this. If the minimum of 96.8% protein sequence identity defines the limit of viral species (Vasilakis *et al.*, 2014), KPhV can never be assigned as a member of the Alphamesonivirus 1/AMV1 species. However, KPhV RdRp shares only 92% sequence identity with those of *bona fide* members of AMV1, and this suggestion is further supported by all the tree topologies obtained. Again, phylogenetic information and distance values clearly indicate KPhV and DkNV (which share 99% of RdRp sequence identity) should be members of the same (AMV3) species, confirming phylogenetic assessments in recent studies (Wang *et al.*, 2017; Newton *et al.*, 2020). In a similar situation, the analysis of the phylogenetic tree topologies clearly showed that BBaV should also not belong to the AMV1 species. Furthermore, when the RdRp sequences of BBaV are compared with those of AMV1 members, as mentioned above for KPhV, their RdRp share only around 90% of sequence identity. Therefore, it should not be classified as a member of AMV1. On the other hand, BBaV does seem to share common ancestry with KSaV, but both these virus' RdRp sequences form independent monophyletic clusters in phylogenetic trees (Fig. 2). While Vasilakis *et al.* (2014) did suggest BBaV and KSaV should be considered as separate species, Wang *et al.* (2017) and Newton *et al.* (2020) place them into the same species (AMV2). While their sequences do share high similarity, their RdRp shared only 96% identity, falling below the 96.8% cut-off value for viral species assignment. Therefore, both the RdRp phylogenetic tree topology and sequence similarity values support previous claim by Vasilakis *et al.* (2014) that these two viruses should be placed into two different viral species in the *Mesoniviridae* family. Since KSaV has been detected first (Vasilakis *et al.*, 2014) and assigned to the AMV2

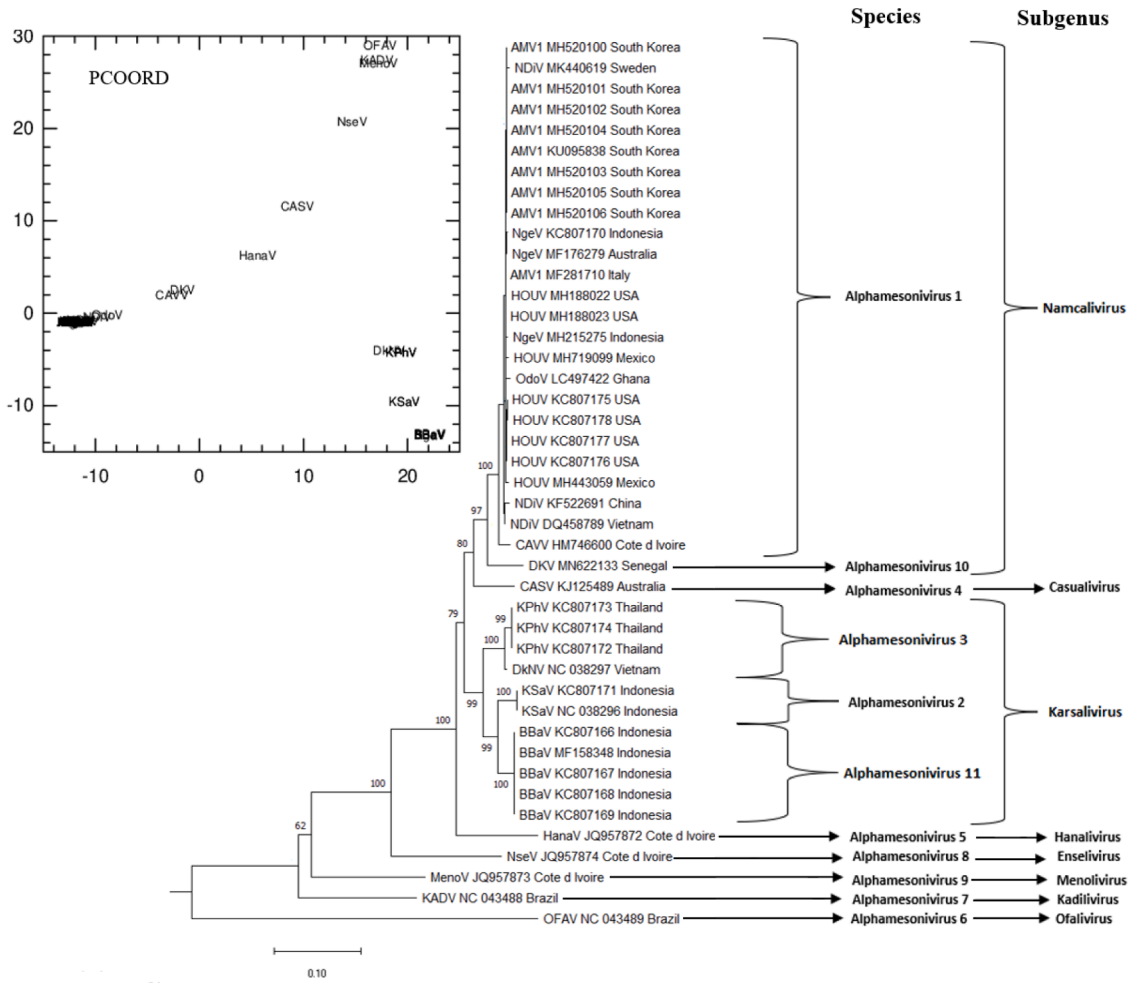

Fig. 2. Maximum likelihood tree for mosquito mesonivirus RdRp protein sequences estimated under a WAG substitution model (right panel), alongside principal coordinate analysis (left panel) carried out using PCOORD, where each sequence is identified by its corresponding abbreviation: HOUV = Houston virus; AMV1 = Alphamesonivirus 1; NDiV = Nam Dinh virus; NgeV = Ngewontan virus; OdoV = Odorna virus; CAVV = Cavally virus; DKV = Dianke virus; HanaV = Hana virus; BBaV = Bontang Baru virus; KSaV = Karang Sari virus; KPhV = Kamphaeng Phet virus; DkNV = Dak Nong virus; CASV = Casuarina virus; NseV = Nse virus; KADV = Kadiweu virus; MenoV = Meno virus; OFAV = Ofaie virus); The species and genera indicated follow the taxonomy revision proposal presented in this work.

species, we tentatively propose that BBaV should, instead, be regarded a member of the new Alphamesonivirus 11 species (Fig. 2). Our analysis also suggested that the Odorna virus from Ghana, which remains unclassified up to the present day, should also belong to the AMV1 sublineage, as it shares over 97% of identity to other AMV1 sequences. The mesonivirus classification at the subgenus level has not yet been extensively studied and should also be reconsidered from what is currently assigned at the NCBI taxonomy browser, where the *Namcalivirus* subgenus not only contains all previously defined mesoniviruses in the AMV1 species, but also KPhV and BBaV, while the *Karsalivirus* only contains DkNV and KSaV. We propose for a reference value of 93% of protein sequence identity (RdRp protein sequences) to be used as a reference for definition of a new subgenus in the *Mesoniviridae* family. DkNV, KPhV, BBaV and KSaV share less than 93% identity between them, and more than 93% against other MM. As such, they should all be inserted into one specific subgenus, in this case the *Karsalivirus* subgenus, while the *Namcalivirus* should only contain the species in its monophyletic clade (seen in Fig. 2, with all its sequences sharing more than 93% identity values). All the remainder sequences, Alphamesonivirus 4 (CASV), Alphamesonivirus 5 (HanaV), Alphamesonivirus 6 (OFAV), Alphamesonivirus 7 (KADV), Alphamesonivirus 8 (NseV), Alphamesonivirus 9 (MenoV) and Alphamesonivirus 10 (DKV), were represented by one single sequence each in the ML phylogenetic tree, where they appear as isolated branches, and their taxonomy classification, both at genus and subgenus level, have been reinforced by the results/findings of this study.

3.4. *Analyses with other mesonivirus and virus from other families of nidoviruses*

To further extend the phylogenetic analyses carried out in this work, we reconstructed the evolutionary relationships of mosquito mesonivirus, not only to other viruses in the *Nidovirales* Order, but also with the recently described mesonivirus identified in non-mosquito hosts. An initial tree was obtained using ORF1ab aa sequences (not shown), but since still no full-length ORF1ab sequence has yet been described for LtM, phylogenetic reconstruction was refined using only RdRp sequences (Supplementary Fig. 7). In the suborder *Mesnidovirineae*, mesoniviruses clearly form a stable clade that shares ancestry with the so-called Beihai Nido-like virus, the single representative of the *Medioniviridae* family, but those found in hosts other than mosquitoes (OM: AcMSV, Immeso1, LtM) are positioned between the large monophyletic clade that defines the mosquito mesonivirus lineage, as independent (not forming a cluster) sister lineages of the latter. The phylogenetic relatedness between MM and OM was not only suggested by the topology of the RdRp tree, but also by the genetic distance values obtained when OM and MM sequences were compared, indicating that OM were consistently closer to MM than to the *Medioniviridae* members (only two

identical sequences have been described, both from the same species, so only one is indicated in Supplementary Fig. 6). Finally, assessments of OM vs OM and OM vs MM protein sequence divergence between OM sequences, led to high divergence values, which further suggests they may correspond to the maiden members of putative new mesonivirus taxon (genera, family). These hypotheses will be investigated as further OM sequences become available.

4. Discussion

In this study, we sought to extend previously published genetic characterization data (Vasilakis *et al.*, 2014) regarding the *Mesoniviridae* family of viruses. Over recent years, the number of mesonivirus sequences deposited in GenBank has significantly increased, which expanded the potential for new genetic analyses and phylogenetic inference analyses. While many newly described sequences clustered into predefined mesonivirus genetic lineages (like most of the sequences of the Alphamesonivirus 1 species), some were classified as totally new species (e.g. the Dianke virus). More importantly, phylogenetic reconstructions and sequence similarity calculations carried out during our study brought out new information that calls for a revision of the classification (taxonomy) of mesoniviruses.

Unlike previous reconstructions (Vasilakis *et al.*, 2014; Chang *et al.*, 2020) which mostly focused on the analysis of S-sequences, we performed analyses based on different sets of nucleotide and protein sequences. While the phylogenetic signals and tree topologies were calculated using multiple sequence datasets corresponding to the total coding-sequence or its ORF-specific fractions (most of which display high phylogenetic signal), special attention was devoted to phylogenetic reconstructions involving the RdRp-specific coding region, which has been extensively used for nidovirus species demarcation criteria (Cowley and Walker, 2014). Focusing on a specific dataset of sequences is important in this specific case, since our data indicate that phylogenetic reconstructions based on different genomic regions (complete-genome or ORF2A/S vs RdRp) does not result in congruent topologies, therefore confounding the establishment of clade demarcation criteria, and consequently complicates the taxonomic classification of these viruses. Therefore, we suggest that mesonivirus taxonomy should focus on the analysis of only RdRp aa sequences, using a minimum reference value of 96.8% of protein sequence identity to define a mesonivirus species, and 93.0% of protein sequence identity as a reference value of to place mesoniviruses into a given subgenus. Not only did our analysis indicate that tree topologies and genetic diversity values at times clash with the prevailing classification scheme, the description of new viral sequences in the coming years will bring new light into the structure of the mesonivirus taxon. We proposed that Odorna (which has remained unclassified up to the present day) should become a member of the Alphamesonivirus 1 species in the *Namcalivirus* subgenus. Also, our analysis suggests that BBaV should be regarded as a candidate for new species in the *Mesoniviridae* family, tentatively named as Alphamesonivirus 11. On the other hand, the *Karsalivirus* genus, currently containing both DkNV (AMV3) and KSaV (AMV2), should also contain both BBaV (AMV11) and KPhV (AMV3). As for the remaining species, our analyses reinforced the currently accepted classification.

While taxonomic decisions based on the analysis of a small section of a viral genome (RdRp), as opposed to the use of the viral genome, may be disputed, RdRp sequences are currently the most frequently represented mesonivirus sequence in the public genomic databases. Future studies should not only focus on the *Mesoniviridae* family but also other nidoviruses and ideally should focus on obtaining full-length sequences, as our results indicate it has the highest phylogenetic signal. These studies should also combine phylogenetic, genetic distance, and statistical analyses (such as PCOORD) as complementary tools for genetic analyses.

Previous observations did indicate a worldwide distribution mesonivirus (Vasilakis *et al.*, 2014), which we further emphasized here, with the analysis of mesonivirus sequences obtained from mosquitos collected in multiple continents. However, no signs were found regarding geographic segregation patterns. There were also no obvious signs of species-specific host restrictions, unlike what has been described for most lineages of insect-specific flaviviruses (Colmant *et al.*, 2017). For example, like what happens to the Alphamesonivirus 1 species, most sequences have been obtained from multiple subspecies of *Culex* or *Aedes* mosquitos from multiple countries. However, the majority of the other Alphamesonivirus species are currently characterized by either one or only a few genetically close members with a similar geographic origin, which confounds the identification of possible geographic or host-range limits.

The detailed analysis of mesonivirus genomic features confirmed that newly described MM sequences conform, in general terms, to the genetic organization previously defined for mesoniviruses. When compared to other nidovirus, mesonivirus have shorter genomes (the only exception being the *Arteriviridae* family, with even shorter genomes) as well as a lower number of identifiable ORFs. The analysis of recently described mesonivirus sequences indicated the presence of a sequence insertion block in ORF1a similar to that reported previously (Vasilakis *et al.*, 2014) in DkNV. Nidovirus genome expansions have been previously reported (Lauber *et al.*, 2013), but the specific functional role of the ORF1a insertion blocks remains unclear. Our analysis suggested it has coding capacity, though its putative product is of unknown function. Other than function of the products of the readily identifiable ORF1a and ORF1b, the putative functions of the other ORFs found in the MM viral sequences analyzed remains elusive. However, remote homology detection and some of their biochemical features suggest two of them encode a nucleocapsid and a viral envelope glycoprotein. Altogether, mesoniviruses are characterized by low amino acid sequence diversity (by entropy assessment), as well as a lower number of non-synonymous substitutions (by calculation of ω values), especially when compared to other nidoviruses.

We executed the first phylogenetic reconstruction with multiple meso-like virus isolated from non-mosquito hosts to elucidate how they fare in phylogenetic relationships into the whole *Nidovirales* Order radiation. They were all classified as members of the *Nidovirales* Order based on its genomic structure, amino acid sequence identity and phylogenetic analysis, expanding our knowledge on the host range of mesonivirus, previously only reported in mosquitos. Although they could tentatively be classified as mesonivirus, their sequence identity with mosquito mesonivirus and virus from other close families, like the *Medioniviridae* and *Coronaviridae* family, is quite low. Even between themselves there is high sequence discrepancy, and there are significant differences in their genomic structure. While the more conserved regions (both ORF1a and ORF1b) look to be nearly identical to other mesonivirus (including its putative proteome characteristics), the remaining ORFs, which should correspond to structural protein coding regions, did not found any similarity searches, with no known function or domain found as well. Further studies are needed, as more non-mosquito meso-like virus are identified in the future, to evaluate whether these new viruses could indeed be clustered with viruses of the *Mesoniviridae* family, or even shape a new family. Also, this does corroborate past studies that hinted at the evolution of nidovirus in arthropods and consequent spread into other group of hosts, including vertebrates (Nga *et al.*, 2011), which may happen again with mesonivirus. Coronavirus from completely different hosts (like bats and equines) share low sequence identity, which also happens between mosquito mesonivirus and other mesonivirus (with non-mosquito hosts).

In contrast with the high phylogenetic signal values associated with of most datasets of mesonivirus sequences, assessment of sequence divergence through time using root-to-tip analysis, systematically indicated, for all datasets used, that there is still insufficient data in the public databases to possibly support a phylogeographic reconstruction of the evolution of mesoniviruses. This result is most probably the consequence of a poor range of sampling time for existing sequences,

which would negatively impact the sequences' temporal signal. In fact, even though some sequences (from BBaV, KPhV and NgeV) were detected in mosquito collections from the early 1980's, the remainder have been mostly obtained from mosquitos collected in the last 10 to 15 years. Therefore, phylogeographic reconstructions that would disclose the geographic distribution of mesoniviruses over time still awaits that more diverse assemblages of heterochronous mesonivirus sequences become available in the near future.

Mesoniviruses look to be an ever-expanding and unique group of viruses in the *Nidovirales* Order, with more information being obtained as new sequences are identified. Even their stance as insect-specific viruses could no longer hold true, as more hosts continue to be recognized (if they indeed end up being classified as virus in the *Mesoniviridae* family). Studies like these should continue to be executed in the future. Their potential to be developed as biological control agents, which have been identified in similar viruses (Goenaga *et al.*, 2020), also remains unclear and is an important area for future investigation.

Supplementary Fig. 1: Schematic representation of nucleotide sequences for each group of mesonivirus and for other virus from different families in the *Nidovirales* Order, with different ORFs identified; * - Nucleotide sequence for the Odorna virus seems incomplete and contains no further information apart from the one present here; RFS: ribosomal frameshift elements; HOUV = Houston virus; AMV1 = Alphamesonivirus 1; NDiV = Nam Dinh virus; NgeV = Ngewontan virus; OdoV = Odorna virus; CAVV = Cavally virus; DKV = Dianke virus; HanaV = Hana virus; BBaV = Bontang Baru virus; KSaV = Karang Sari virus; KPhV = Kamphaeng Phet virus; DkNV = Dak Nong virus; CASV = Casuarina virus; NseV = Nse virus; KADV = Kadiweu virus; MenoV = Meno virus; OFAV = Ofaie virus; SARS-CoV-2 = Severe acute respiratory syndrome coronavirus 2 (*Coronaviridae*; MT997203); SheV = Simian hemorrhagic encephalitis virus (*Arteriviridae*; NC_038293); FmN = Fathead minnow nidovirus (*Tobaniviridae;* NC_038295); BlN = Botrylloides leachii nidovirus (*Medionivirineae*; MK956105).

Supplementary Fig. 2: Heat map representing intersequence genetic diversity of mesonivirus. Representative tree (maximum likelihood, WAG model) based on RdRp aa sequences (sequences identifiable in Supplementary Table 1), and Z-Scores were obtained based on pairwise evolutionary distances obtained on MegaX.

Supplementary Fig. 3: Intragroup genetic diversity of mesonivirus. Representative RdRp tree (maximum likelihood, WAG model) based on the analysis of alignments of RdRp primary sequences. For each species, sequence identification follows the nomenclature indicated in Supplementary Table 1, followed by number of sequences for each clade; box-and-whisker graphs are used to plot distributions of pairwise evolutionary distances of three different sets: between mesoniviruses from the same species (Alphamesonivirus 1/AMV1), between all mosquito mesoniviruses (MM), and between all mosquito mesoniviruses (MM) and other mesoniviruses (OM).

Supplementary Fig. 4: Entropy calculations based the Shannon function (Shannon entropy-one) applied on alignments of ORF1a protein sequences from different families in the *Nidovirales* Order.

Supplementary Fig. 5: Principal coordinate analysis carried (left panel) out for mosquito mesonivirus S protein coding sequences. Each sequence is identified by the sequence abbreviation they belong to (HOUV = Houston virus; AMV1 = Alphamesonivirus 1; NDiV = Nam Dinh virus; NgeV = Ngewontan virus; OdoV = Odorna virus; CAVV = Cavally virus; DKV = Dianke virus; HanaV = Hana virus; BBaV = Bontang Baru virus; KSaV = Karang Sari virus; KPhV = Kamphaeng Phet virus; DkNV = Dak Nong virus; CASV = Casuarina virus; NseV = Nse virus; KADV = Kadiweu virus; MenoV = Meno virus; OFAV = Ofaie virus). A maximum likelihood tree (right panel), estimated under a WAG substitution model, is also shown, while displaying the taxonomy revision proposal presented in this work.

Supplementary Fig. 6: Maximum likelihood tree for mosquito mesonivirus full-length sequences. Each sequence is identified by the sequence abbreviation they belong to (HOUV = Houston virus; AMV1 = Alphamesonivirus 1; NDiV = Nam Dinh virus; NgeV = Ngewontan virus; OdoV = Odorna virus; CAVV = Cavally virus; DKV = Dianke virus; HanaV = Hana virus; BBaV = Bontang Baru virus; KSaV = Karang Sari virus; KPhV = Kamphaeng Phet virus; DkNV = Dak Nong virus; CASV = Casuarina virus; NseV = Nse virus; KADV = Kadiweu virus; MenoV = Meno virus; OFAV = Ofaie virus).

Supplementary Fig. 7: Maximum likelihood tree for protein sequences (RdRp coding region) of virus from different families in the *Nidovirales* Order.

CRediT authorship contribution statement

Paulo Morais: Data curation, Formal analysis, Visualization. **Nídia S. Trovão:** Writing – review & editing. **Ana B. Abecasis:** Writing – review & editing. **Ricardo Parreira:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work received financial support from the Global Health and Tropical Medicine Center, which is funded through FCT contract UID/Multi/04413/2013. The opinions expressed in this article are those of the authors and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.virusres.2022.198727.

References

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S., 2016. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 44 (W1), W147–W153. https://doi.org/10.1093/nar/gkw419.

Blitvich, B.J., Firth, A.E., 2015. Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization. Viruses 7 (4), 1927–1959. https://doi.org/10.3390/v7041927.

Boni, M.F., Lemey, P., Jiang, X., Lam, T.T.Y., Perry, B.W., Castoe, T.A., Robertson, D.L., 2020. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. Nat. Microbiol. 5 (11), 1408–1417. https://doi.org/10.1038/s41564-020-0771-4.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17 (4), 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

Chang, T., Guo, M., Zhang, W., Niu, J., Wang, J.J., 2020. First report of a mesonivirus and its derived small RNAs in an aphid species aphis citricidus (Hemiptera: Aphididae), implying viral infection activity. J. Insect Sci. 20 (2), 1–4. https://doi.org/10.1093/jisesa/ieaa022.

Charles, J., Tangudu, C.S., Hurt, S.L., Tumescheit, C., Firth, A.E., Garcia-Rejon, J.E., Blitvich, B.J., 2018. Detection of novel and recognized RNA viruses in mosquitoes from the Yucatan Peninsula of Mexico using metagenomics and characterization of their *in vitro* host ranges. J. Gen. Virol. 99 (12), 1729–1738. https://doi.org/10.1099/jgv.0.001165.

Chiapello, M., Bosco, L., Ciuffo, M., Ottati, S., Salem, N., Rosa, C., Turina, M., 2021. Complexity and local specificity of the virome associated with tospovirus-transmitting thrips species. J. Virol. 95 (21) https://doi.org/10.1128/jvi.00597-21.

Colmant, A.M.G., Hobson-Peters, J., Bielefeldt-Ohmann, H., van den Hurk, A.F., Hall-Mendelin, S., Chow, W.K., Hall, R.A., 2017. A new clade of insect-specific flaviviruses from Australian anopheles mosquitoes displays species-specific host restriction. MSphere 2 (4), 155–166. https://doi.org/10.1128/mSphere.00262-17.

Cowley, J.A., Walker, P.J., 2014. Molecular biology and pathogenesis of roniviruses. Nidoviruses 361–377. https://doi.org/10.1128/9781555815790.ch24.

Diagne, M.M., Gaye, A., Ndione, M.H.D., Faye, M., Fall, G., Dieng, I., Sall, A.A., 2020. Dianke virus: a new mesonivirus species isolated from mosquitoes in eastern

Senegal. Virus Res. 275 (June 2019), 197802 https://doi.org/10.1016/j.virusres.2019.197802.
Goenaga, S., Goenaga, J., Boaglio, E.R., Enria, D.A., Levis, S.D.C., 2020. Superinfection exclusion studies using west nile virus and culex flavivirus strains from Argentina. Mem. Inst. Oswaldo Cruz. 115 (5), 1–5. https://doi.org/10.1590/0074-02760200012.
Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., Snijder, E.J., 2006. Nidovirales: evolving the largest RNA virus genome. Virus Research 117, 17–37. https://doi.org/10.1016/j.virusres.2006.01.017.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95-98-NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hon, C.C., Lam, T.Y., Shi, Z.L., Drummond, A.J., Yip, C.W., Zeng, F., Leung, F.C.C, 2008. Evidence of the recombinant origin of a bat severe acute respiratory syndrome (SARS)-like coronavirus and its implications on the direct ancestor of SARS coronavirus. J. Virol. 82 (4), 1819–1826. https://doi.org/10.1128/JVI.01926-07.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30 (14), 772–780. https://doi.org/10.1093/molbev/mst010.
Kosakovsky Pond, S.L., Frost, S.D.W, 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21 (10), 2531–2533. https://doi.org/10.1093/bioinformatics/bti320.
Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547–1549. https://doi.org/10.1093/molbev/msy096.
Lai, C., Shih, T., Ko, W., Tang, H., Hsueh, P., 2020. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. Int. J. Antimicrob. Agents 55 (3), 105924. https://doi.org/10.1016/j.ijantimicag.2020.105924.
Lauber, C., Goeman, J.J., de Parquet, M.C., Thi Nga, P., Snijder, E.J., Morita, K., Gorbalenya, A.E., 2013. The footprint of genome architecture in the largest genome expansion in RNA viruses. PLoS Pathog. 9 (7) https://doi.org/10.1371/journal.ppat.1003500.
Letunic, I., Bork, P., 2018. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 46 (D1), D493–D496. https://doi.org/10.1093/nar/gkx922.
Marklewitz, M., Zirkel, F., Rwego, I.B., Heidemann, H., Trippner, P., Kurth, A., Junglen, S., 2013. Discovery of a unique novel clade of mosquito-associated bunyaviruses. J. Virol. 87 (23), 12850–12865. https://doi.org/10.1128/jvi.01862-13.
Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1 (1), 1–5. https://doi.org/10.1093/ve/vev003.
Nei, M., Gojoborit, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426. https://doi.org/10.1093/oxfordjournals.molbev.a040410.
Newton, N.D., Colmant, A.M.G., O'Brien, C.A., Ledger, E., Paramitha, D., Bielefeldt-Ohmann, H., Hobson-Peters, J., 2020. Genetic, morphological and antigenic relationships between mesonivirus isolates from Australian mosquitoes and evidence for their horizontal transmission. Viruses 12 (10). https://doi.org/10.3390/v12101159.
Nga, P.T., de Parquet, M.C., Lauber, C., Parida, M., Nabeshima, T., Yu, F., Gorbalenya, A. E., 2011. Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. PLoS Pathog. 7 (9) https://doi.org/10.1371/journal.ppat.1002215.
Rambaut, A., Lam, T.T., Carvalho, L.M., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2 (1) https://doi.org/10.1093/ve/vew007.
Schoch, C.L., Ciufo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Karsch-Mizrachi I. et al. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. Database, 2020. 10.1093/database/baaa062.
Schmidt, H.A., Strimmer, K., Vingron, M., Von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18 (3), 502–504. https://doi.org/10.1093/bioinformatics/18.3.502.
Strimmer, K., Von Haeseler, A., 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Nat. Acad. Sci. USA 94 (13), 6815–6819. https://doi.org/10.1073/pnas.94.13.6815.
Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44 (W1), W232–W235. https://doi.org/10.1093/nar/gkw256.
Vasilakis, N., Guzman, H., Firth, C., Forrester, N.L., Widen, S.G., Wood, T.G., Tesh, R.B., 2014. Mesoniviruses are mosquito-specific viruses with extensive geographic distribution and host range. Virol. J. 11 (1), 1–12. https://doi.org/10.1186/1743-422X-11-97.
Wang, Y., Xia, H., Zhang, B., Liu, X., Yuan, Z., 2017. Isolation and characterization of a novel mesonivirus from Culex mosquitoes in China. Virus Res. 240 (June), 130–139. https://doi.org/10.1016/j.virusres.2017.08.001.
Warrilow, D., Watterson, D., Hall, R.A., Davis, S.S., Weir, R., Kurucz, N., Hobson-Peters, J., 2014. A new species of mesonivirus from the northern territory, Australia. PLoS ONE 9 (3). https://doi.org/10.1371/journal.pone.0091103.
Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kübler, J., Lozajic, M., Alva, V., 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol. 430 (15), 2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007.
Zirkel, F., Kurth, A., Quan, P.L., Briese, T., Ellerbrok, H., Pauli, G., Junglen, S., 2011. An insect nidovirus emerging from a primary tropical rainforest. MBio 2 (3), 21–23. https://doi.org/10.1128/mBio.00077-11.