Identification and Characterization of Genetically Divergent Members of the Newly Established Family *Mesoniviridae*

Florian Zirkel,^a Hanna Roth,^a Andreas Kurth,^b Christian Drosten,^a John Ziebuhr,^c Sandra Junglen^a

Institute of Virology, University of Bonn Medical Center, Bonn, Germany^a; Center for Biological Safety, Robert Koch Institute, Berlin, Germany^b; Institute of Medical Virology, Justus Liebig University, Giessen, Germany^c

The recently established family *Mesoniviridae* (order *Nidovirales*) contains a single species represented by two closely related viruses, Cavally virus (CavV) and Nam Dinh virus (NDiV), which were isolated from mosquitoes collected in Côte d'Ivoire and Vietnam, respectively. They represent the first nidoviruses to be discovered in insects. Here, we report the molecular characterization of four novel mesoniviruses, Hana virus, Méno virus, Nsé virus, and Moumo virus, all of which were identified in a geographical region in Côte d'Ivoire with high CavV prevalence. The viruses were found with prevalences between 0.5 and 2.8%, and genome sequence analyses and phylogenetic studies suggest that they represent at least three novel species. Electron microscopy revealed prominent club-shaped surface projections protruding from spherical, enveloped virions of about 120 nm. Northern blot data show that the four mesoniviruses analyzed in this study produce two major 3'-coterminal subgenomic mRNAs containing two types of 5' leader sequences resulting from the use of different pairs of leader and body transcription-regulating sequences that are conserved among mesoniviruses. Protein sequencing, mass spectroscopy, and Western blot data show that mesonivirus particles contain eight major structural protein species, including the putative nucleocapsid protein (25 kDa), differentially glycosylated forms of the putative membrane protein (20, 19, 18, and 17 kDa), and the putative spike (S) protein (77 kDa), which is proteolytically cleaved at a conserved site to produce S protein subunits of 23 and 57 kDa. The data provide fundamental new insight into common and distinguishing biological properties of members of this newly identified virus family.

During a surveillance study on the impact of ecological changes on mosquito-borne pathogens in Côte d'Ivoire, a virus with distant relationship to coronaviruses based on a fragment from the putative RNA-dependent RNA polymerase (RdRp) gene was discovered (1). The tentatively named Cavally virus (CavV) was subsequently shown to represent the first insect-associated nidovirus, defining a novel virus family within the order *Nidovirales* (2). Independently, a virus, termed Nam Dinh virus (NDiV), was identified in mosquitoes collected in Vietnam and also proposed as the prototype of a novel virus family in the *Nidovirales* (3). The viruses turned out to be closely related (4) and are now proposed to belong to the same virus species, termed *Alphamesonivirus 1*, founding the solitary genus *Alphamesonivirus* within the new virus family *Mesoniviridae*. Pending final International Committee on Taxonomy of Viruses (ICTV) ratification (http://talk.ictvonline.org/files/proposals/taxonomy_proposals_invertebrate1/m/inv04/4378.aspx), the order *Nidovirales* will then contain four families: *Arteriviridae*, *Coronaviridae* (subfamilies *Coronavirinae* and *Torovirinae*), *Mesoniviridae*, and *Roniviridae*.

In spite of their close genetic relatedness, different particle morphologies were described for CavV and NDiV. Mature CavV virions were shown to appear as enveloped spherical particles 120 nm in diameter with large club-shaped surface projections (2), while immature CavV particles detected in infected cells were 50 to 60 nm in diameter and lacked surface projections (2). In contrast, NDiV was reported to be an enveloped, spherical virus that seemed to lack surface projections and had a size of about 60 to 80 nm (3).

The positive-sense RNA genomes of CavV and NDiV comprise 20,187 and 20,192 nucleotides (nt), respectively, placing these viruses, in terms of genome size, between the large nidoviruses (*Coronavirus* [CoV], *Torovirus* [ToV], and *Ronivirus* [RoV], 26 to 32 kb) on one side and the small nidoviruses (*Arterivirus* [ArV], 13

to 16 kb) on the other. The medium-size genomes of the two viruses, CavV and NDiV, inspired the name mesonivirus (in Greek, *mesos* means “in the middle,” while “ni” refers to nidoviruses [4]). *Alphamesonivirus 1* genomes were consistently reported to contain seven major open reading frames (ORFs), with ORF1a and -1b located at the 5' end, encompassing two-thirds of the genome, and the smaller ORF2a, -2b, -3a, -3b, and -4 occupying the 3'-proximal one-third of the genome. Similar to other nidoviruses (reviewed in reference 5), ORF1a and -1b overlap by a few nucleotides and are predicted to encode two replicase polyproteins (pp) called pp1a and pp1ab (2, 3). Expression of the replicase polyprotein is probably controlled by ribosomal frameshifting (RFS), as shown for other nidoviruses (for a recent review, see reference 6). The conserved heptanucleotide motif GGAUUUU has been proposed to act as a “slippery sequence” during frameshifting (2, 3). Conserved domains in pp1ab of mesoniviruses and large nidoviruses (CoV, ToV, and RoV) include (from the N to the C terminus) a picornavirus 3C-like main protease (3CL^{Pro}, also called M^{Pro}), which is flanked on either side by two transmembrane domains (TMDs), an RdRp, a complex zinc-binding motif (Z) that is linked to a superfamily 1 helicase (Hel), a 3'-to-5' exoribonuclease (ExoN), a guanine-N7 methyltransferase (NMT), and a ribose-2'-O-methyltransferase (OMT) (2–4). In contrast to

Received 12 February 2013 Accepted 18 March 2013

Published ahead of print 27 March 2013

Address correspondence to Sandra Junglen, junglen@virology-bonn.de.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/JVI.00416-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.00416-13

nidoviruses infecting vertebrate hosts (ArV, CoV, and ToV), the invertebrate-infecting mesoniviruses do not appear to encode a uridylylate-specific endonuclease (NendoU) that previously was thought to be a nidovirus-wide conserved domain (3, 4, 7–9). Also, the conservation of two other domains, ExoN and OMT, varies among different lineages of nidoviruses (3, 7, 9).

The characterization of viral RNA synthesis in CavV- and NDiV-infected cells suggested that, like other nidoviruses, mesoniviruses express the genes downstream of the replicase gene from a nested set of 3′-coterminal subgenomic (sg) mRNAs (2, 3). CavV was further shown to generate sg mRNAs that are 3′ and 5′ coterminal and carry 5′ leader sequences derived from the 5′ end of the genome, as was shown previously for most other nidoviruses (2, 10, 11). Two sg mRNAs were mapped to ORF2a/b and ORF3a/b by Northern blotting (2). As a unique mechanism among nidoviruses, synthesis of mesonivirus sg mRNAs seems to involve two different leader transcription-regulating sequence (TRS) elements, resulting in mRNA species with 5′ leaders of varying lengths (2, 3). The positions and sizes of TRSs, as well as the number of sg mRNAs detected in virus-infected cells, varied between the two studies.

The five major ORFs in the 3′-terminal genome regions of CavV and NDiV were predicted to encode a spike (S) glycoprotein (ORF2a), a nucleocapsid (N) protein (ORF2b), two proteins with membrane-spanning regions (ORF3a and -3b), and a small protein with unknown function (ORF4) (2, 3). Three virion proteins encoded by ORF2a, -2b, and -3a were identified for NDiV by SDS-PAGE and Edman sequencing (3).

In this study, we sought to gain further insight into the prevalence and genetic diversity of mesoniviruses in our sample of mosquitoes from Côte d'Ivoire. We report the discovery, sequence analysis, and characterization of four novel mesoniviruses that, as was previously done for CavV, were named after rivers in the Tai National Park. Furthermore, we performed experiments focusing on virion morphology, transcription strategy, and viral structural proteins, providing interesting new insight into common and distinguishing biological properties of members of this newly identified nidovirus lineage.

MATERIALS AND METHODS

Virus isolation, propagation, and purification. Mosquitoes were collected in the area of the Tai National Park in Côte d'Ivoire (1). Virus isolation experiments were done from 432 pools with 4,839 female mosquitoes on C6/36 (*Aedes albopictus*) and Vero cells as described previously (1, 12). Mosquito pools (1 to 50 specimens) were generated according to species, sex, and sampling location. Mosquitoes that were rarely found were combined with other rare species from the same sampling location. To obtain pure virus stocks from pools infected with multiple viruses, two endpoint dilutions via titration on C6/36 cells were performed using infectious cell culture supernatant from the 2nd passage (isolates CavV/C79/CI/2004, Hana virus/A4/CI/2004 [HanaV], Nsé virus/F7/CI/2004 [NséV], and Méno virus/E9/CI/2004 [MénoV]) (2). Viral replication was measured by real-time PCR with virus-specific primers and probes (Hana virus F, 5′-ATCATAACTGATGGCATGAAAGATG-3′; Hana virus R, 5′-CAATGCGGTCTGTGTCGAT-3′; Hana virus TM, 5′-6-carboxyfluorescein [FAM]-TTAAAGCCACACCCTTACATCACCTACCGC-black hole quencher 1 (BHQ1)-3′; Nsé virus F, 5′-AATTATAACTGAGGGAA TGAAGATGAAC-3′; Nsé virus R, 5′-CAAATGCAATCCTGATCAAG CA-3′; Nsé virus TM, 5′-FAM-AAACCTCACCCCTATATAACAC-CCCGCACT-BHQ1-3′; Méno virus F, 5′-TGGAACCAACAATTGCCG ATA-3′; Méno virus R, 5′-ACATTGGCTGAGATGGGATT-3′; Méno virus TM, 5′-FAM-CGCTCGTACATTCACCGTTTACCC-BHQ1-3′).

Virus growth kinetics. C6/36 cells were infected with virus stocks from the 4th passage with a multiplicity of infection (MOI) of 0.001 as described previously (2). An aliquot of cell culture supernatant was removed 0, 6, 12, 24, and 48 h postinfection (p.i.). RNA was extracted using the NucleoSpin RNA virus kit (Machery & Nagel, Dueren, Germany), and cDNA was synthesized using the SuperScript III RT system (Life Technologies, Karlsruhe, Germany). Virus replication was measured by real-time PCR.

Infection of vertebrate cells. To investigate replication in vertebrate cells, VeroE6 and VeroFM cells (African green monkey kidney cells; ATCC CRL-1586 and CCL-81) were inoculated with CavV, HanaV, NséV, and MénoV and incubated for 7 days. Five blind passages of cell culture supernatant were performed on fresh cells in 1:10 dilutions. Furthermore, three blind passages of CavV and NséV were performed on Caco-2 cells (human colon carcinoma cells; ATCC HTB-37), HEF (human embryonic fibroblasts; ATCC CRL-7093), RAW 264.7 cells (mouse leukemic monocyte macrophages; ATCC TIB-71), FeA cells (feline embryo cells, kindly provided by Marcel Asper, NewLab Inc., Cologne, Germany), PSEK cells (porcine stable equine kidney cells, kindly provided by Thomas Briese, Columbia University, New York, NY), and RoNi cells (*Rousettus aegyptiacus* kidney cells [13]). Virus replication was measured by real-time PCR.

Electron microscopy. Cell culture supernatants and ultrathin sections of infected C6/36 cells were analyzed by electron microscopy as described previously (2).

PCR screening. RNA was extracted from homogenized mosquito pools ($n = 432$) and from pools inducing cytopathic effects (CPE) on C6/36 cells ($n = 105$) using the QIAmp Viral RNA Mini Kit (Qiagen, Hilden, Germany). cDNA was synthesized using the Superscript III RT System (Life Technologies, Karlsruhe, Germany) with random hexameric primers (R6) following the manufacturer's instructions. Infectious cell culture supernatants were tested by nested PCR using Phusion DNA polymerase (Thermo Scientific, Braunschweig, Germany) and the primers MeNiV-F1 (5′-AAAYAAATCAGAASCAGGACG-3′) and MeNiV-R1 (5′-WGTYACRCCTCCDGGTTTCTG-3′) for first-round PCR and MeNiV-F2 (5′-GCACAATAYGGCGGTTGG-3′) and MeNiV-R2 (5′-RAC TTGHGTRGTTTCAGCCAT-3′). Mosquito homogenates were screened by seminested and fully nested PCRs using Platinum Taq DNA polymerase (Life Technologies, Karlsruhe, Germany) and the primer pairs MeNiV-F3 (5′-CCCATGGTCCCTTCAGAAGT-3′) and MeNiV-R3 (5′-GAAGTTA CACCTCCAGGTTTCTGA-3′) for first-round PCR and combinations of MeNiV-F4 (5′-GGTGATTTCAGAATTCATGCGT-3′), MeNiV-R4 (5′-G TTTTCAGCCATRTATTTCATGCCA-3′), and MeNiV-R5 (5′-CCAACCG CCATATTGTGC-3′) for additional nested PCRs. PCR products were visualized by agarose gel electrophoresis and purified either by gel extraction (QIAex II; Qiagen, Hilden, Germany) or by enzymatic purification (ExoSAP-it, High Wycombe, United Kingdom). Sequencing on both strands was performed by SeqLab (Göttingen, Germany).

Genome sequencing. For full genome sequencing, fragments were elongated by primer walking, seminested PCRs using Platinum Taq DNA polymerase (Life Technologies, Karlsruhe, Germany), and serial fragment-specific primers and primers based on the CavV genome. Genome termini were amplified by 3′ and 5′ rapid amplification of cDNA ends (RACE) (Life Technologies, Karlsruhe, Germany). RACE PCR products were cloned into the TopoTA pCR4 cloning vector (Life Technologies, Karlsruhe, Germany) and amplified. PCR products were analyzed and sequenced as described above.

Genome analysis. Genome sequences were analyzed using Geneious 6 (14). Pairwise identities on the nucleotide and amino acid levels were determined using ClustalW within Geneious 6. Potential signal peptides, transmembrane domains, and N-linked and O-linked glycosylation sites were identified using SignalP4.0 (<http://www.cbs.dtu.dk/services/SignalP/>), TMHMM v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>), and NetCGlyc (<http://www.cbs.dtu.dk/services/NetCGlyc/>), respectively. Phosphorylation sites were identified in NetPhos 2.0 (<http://www.cbs.dtu.dk/services/NetPhos/>). The molecular weights and isoelectric points of

putative proteins were calculated with the pI/MW tool (http://web.expasy.org/compute_pi/).

Phylogenetic analysis. Amino acid sequences were aligned using MAFFT (15) and the E-INS-I algorithm in Geneious 6 (14). The alignment of mesoniviruses with representative members of the order *Nidovirales* was focused on conserved regions. Phylogenetic analyses were conducted using PHYML (16) with the BLOSUM62 amino acid substitution model and rate heterogeneity among sites (4 categories) with 1,000 bootstrap replicates in Geneious 6.

Identification of subgenomic mRNAs, leader-body fusion sites, and TRS elements. Northern blot analyses of viral RNAs were done as described previously (2). The total RNA was further used to amplify the 5' ends of subgenomic mRNAs with gene-specific primers for ORF2a/b, ORF3a/b, and ORF4 using the GeneRacer kit (Life Technologies, Karlsruhe, Germany) according to the manufacturer's instructions. Sequences were analyzed with Geneious 6 software (14).

Protein analysis. Viruses were purified by gradient ultracentrifugation in a sucrose gradient (2). Viral particles were either lysed directly in 1× Laemmli buffer (17) for 10 min at 94°C or deglycosylated using PNGase F (New England BioLabs, Ipswich, MA) following the manufacturer's instructions and subsequently lysed. The protein concentration was measured by Bradford assay using bovine serum albumin as a standard. Equal protein amounts were separated by SDS-PAGE on a 12% or 14% polyacrylamide gel. The gel was washed twice in distilled water and stained with Coomassie (40% methanol, 10% acetic acid, 0.25% Coomassie brilliant blue R-250) for 1 h at 50°C. Multiple washing steps with fixing solution (40% methanol, 10% acetic acid) were performed until excessive Coomassie stain was removed and distinct bands were visible. The bands were analyzed by matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometry. Data were compared to translated mesonivirus ORFs using Geneious 6 (14). N-terminal amino acid sequencing was performed by Edman degradation. Proteins to be analyzed were separated by SDS-PAGE and transferred onto a SequiBlot polyvinylidene difluoride (PVDF) membrane (Bio-Rad, Munich, Germany) according to the manufacturer's instructions. Following electrophoretic transfer, proteins were stained for 10 min in 40% methanol containing 0.025% Coomassie brilliant blue R-250. The membrane was then washed with 50% methanol until distinct protein bands became visible, and individual protein bands were isolated and analyzed by Edman degradation (G. Lochnit, Biochemical Institute, University of Giessen, Giessen, Germany).

Antibody production. CavV and NsÉV were concentrated by ultracentrifugation and inactivated with β-propiolactone (18). Inactivated virus produced from 1×10^{10} 50% tissue culture infective doses (TCID₅₀) of CavV and NsÉV (50% each) was mixed with Freund's complete adjuvant and used to subcutaneously immunize a New Zealand White rabbit three times (days 0, 14, and 28) (Thermo Scientific, Rockford, IL). Antisera (rabbit anti-CavV/NsÉV) were collected at days 42 and 66. In addition, epitope-specific antisera (for details, see Table 3) were raised against each putative protein of CavV (2). The respective antisera were generated by Eurogentec Inc. (Seraing, Belgium; 28-day speedy polyclonal packages).

Western blotting. SDS-PAGE with purified, concentrated viral particles was performed as described above. Proteins were transferred onto a nitrocellulose membrane (Whatman Protran; Sigma-Aldrich, St. Louis, MO) using a semidry blotting system. Transfer efficiency was confirmed by Fast Green FCF staining (Sigma-Aldrich, St. Louis, MO) and subsequent washing of the membrane with 10% acetic acid for 10 min. Non-specific antibody binding was reduced by preincubation of the membrane with 10% RotiBlock solution (Carl Roth, Karlsruhe, Germany). The membrane was incubated with rabbit anti-CavV/NsÉV serum in a 1:1,000 dilution in phosphate-buffered saline (PBS) supplemented with 0.05% Tween 20 solution. Bound antibodies were visualized on an X-ray film using goat anti-rabbit-peroxidase dismutase (POD) (ImmunoPure Antibodies, Thermo Scientific, Braunschweig, Germany) and SuperSignal

West Pico Chemiluminescent substrate (Thermo Scientific, Braunschweig, Germany).

Nucleotide sequence accession numbers. The complete genome sequences of HanaV/A4/CI/2004, MénoV/E9/CI/2004, and NsÉV/F24/CI/2004 were assigned GenBank accession numbers JQ957872 to -74, respectively. The MoumoV sequence fragment of 7,985 nt was assigned GenBank accession number KC768950.

RESULTS

Isolation and detection rates of four novel mesoniviruses. In a previous study, a short genome fragment of a putative further mesonivirus species (A4/CI/2004) with a nucleotide sequence identity of 85% to 86.5% in the RdRp gene compared to CavV was detected (2). The virus was tentatively named Hana virus (isolate HanaV/A4/CI/2004). To test for further mesoniviruses, cell culture supernatants of homogenized mosquito pools that induced CPE in C6/36 cells ($n = 105$) (1) were tested by reverse transcription (RT)-PCR. Besides the detection of CavV-specific sequences in 39 pools (1, 2), CavV-related sequences with numerous double peaks in the sequence chromatograms were identified in 28 out of 105 pools. Individual plasmid clones obtained from the RT-PCR products yielded short genome fragments of CavV and of two distinct viruses with 74.8 to 85.4% nucleotide sequence identity, respectively, in the their RdRp genes (compared to CavV). The two viruses were tentatively named NsÉ virus (isolate NsÉV/F24/CI/2004) and Méno virus (isolate MénoV/E9/CI/2004). To detect viruses that do not efficiently replicate in cell culture, a total of 432 pools of homogenized mosquitoes that were used for the cell culture experiments described above were tested by generic mesonivirus-specific nested PCR. In two pools, a fourth virus with 74.1 to 82.4% nucleotide identity with CavV was identified. The virus was named Moumo virus (MoumoV). Virus host and habitat associations, as well as the prevalences of the respective viruses, are summarized in Table 1.

Mesonivirus morphology and growth. To compare the morphologies and growth characteristics of the newly identified mesonivirus isolates, individual virus clones were purified by endpoint titration. In the case of MénoV, virus isolation attempts from pools MénoV/E12/CI/2004 and MénoV/E13/CI/2004 failed. The only mosquito pool from which MénoV could be isolated (MénoV/E9/CI/2004) proved to be coinfecting with CavV. A pure virus stock of MénoV/E9/CI/2004 was obtained by endpoint titration. Multiple attempts to isolate MoumoV/C88/CI/2004 were not successful.

CavV/C79/CI/2004, HanaV/A4/CI/2004, MénoV/E9/CI/2004, and NsÉV/F7/CI/2004 replicated to high titers of 5×10^9 , 2×10^9 , 2×10^8 , and 7×10^8 TCID₅₀/ml, respectively, in C6/36 cells. HanaV, MénoV, and NsÉV induced only slight morphological changes, while CavV caused a pronounced CPE. Viral RNA in supernatants was measured by real-time PCR over a period of 48 h. The amount of viral RNA increased by up to 10^{10} RNA copies/ml in the first 24 h and peaked at up to 10^{11} RNA copies/ml by 48 h p.i. (Fig. 1A).

To study the permissiveness of vertebrate cells for mesoniviruses, CavV, HanaV, MénoV, and NsÉV were blind passaged five times in VeroE6 and VeroFM cells. CavV and NsÉV were additionally passaged three times in HEF and Caco-2, RAW 264.7, FeA, PSEK, and RoNi cells. No evidence for replication was obtained by real-time RT-PCR for any virus in any of the cells.

Electron microscopy revealed similar morphologies for all the

TABLE 1 Mosquito pools infected with CavV, HanaV, NséV, MénoV, and MoumoV

Sampling site	Pool	Mosquito species ^a	No. of mosquitoes	Virus
Camps	A4	<i>Culex</i> spp.	20	HanaV
	A5	<i>Culex</i> spp.	16	HanaV
	A26	<i>Culex nebulosus</i>	10	NséV, CavV
	A28	<i>C. nebulosus</i>	22	NséV, CavV
	A29	<i>Culex decens</i>	1	CavV ^b
	A61	<i>Culex quinquefasciatus</i>	1	CavV ^b
Primary forest	B42	<i>C. nebulosus</i> , <i>Culex cinerellus</i>	9	NséV, CavV
	B80	ND	24	CavV ^b
Secondary forest	C01	<i>Culex</i> spp.	11	CavV ^b
	C42	<i>C. nebulosus</i>	20	NséV, CavV
	C44	<i>C. nebulosus</i>	20	NséV ^b , CavV ^b
	C45	<i>C. nebulosus</i>	16	NséV, CavV
	C84	<i>Aedes</i> spp.	12	NséV ^b
	C88	ND	20	MoumoV ^b
Plantations	D28	<i>Anopheles</i> spp.	2	NséV, CavV
	D30	<i>Culex</i> spp.	21	NséV, CavV
	E9	<i>Uranotaenia chorleyi</i>	7	MénoV, CavV
	E11	ND	6	MénoV ^b
	E12	<i>Culex</i> spp.	4	MénoV ^b
Human settlements	F7	<i>Anopheles</i> spp.	19	NséV
	F8	<i>Anopheles</i> spp.	30	NséV, CavV
	F24	<i>C. nebulosus</i>	30	NséV, CavV, MoumoV ^b
	F29	<i>C. nebulosus</i>	20	CavV ^b
	F37	<i>C. quinquefasciatus</i>	2	CavV ^b
	F40	<i>Uranotaenia mashaensis</i>	10	CavV ^b
	F52	<i>C. quinquefasciatus</i>	20	CavV ^b

^a ND, not determined.^b No virus isolate available.

viruses included in this study. Viral particles had spherical shapes with mean diameters of 120 nm and carried spikes protruding from their envelopes by about 12 nm (Fig. 1B).

Genome organization of the novel mesoniviruses. Complete genome sequences of HanaV/A4/CI/2004, MénoV/E9/CI/2004, and NséV/F24/CI/2004 were determined using viral RNA isolated from supernatants of infected cells. Because of the limited amount of original sample available for MoumoV/C88/CI/2004, only a sequence fragment of 7,985 nt covering almost the entire ORF1b and parts of the ORF2a and -2b of this noncultivable virus was analyzed. The genomes of HanaV/A4/CI/2004, MénoV/E9/CI/2004, and NséV/F24/CI/2004 were found to comprise 20,070 nt, 19,979 nt, and 20,074 nt [excluding the poly(A) tail], respectively.

The genome sizes and genomic organizations of the newly identified mesoniviruses were similar to those reported for CavV and NDiV (2, 3). 5' untranslated region (UTR) sizes ranged from 290 to 362 nt. The six major ORFs, ORF1a, -1b, -2a, -2b, -3a, and -3b, and the replicase domains TM1-TM2-3CL^{pro}-TM3-RFS-RdRp-Z-Hel-ExoN-NMT-OMT identified in CavV and NDiV (2, 3) were also found to be conserved in HanaV, MénoV, and NséV

(Fig. 2). Alignments of representative conserved (sub)domains of the replicase polyproteins 1a and 1ab (3CL^{pro}, Z, Hel, ExoN, NMT, and OMT) of mesoniviruses are provided in Fig. S1 in the supplemental material. Nucleotide and amino acid identities of mesonivirus replicase domains and ORFs ranged between 58.5 and 89.2%, and 47.5 and 85.3%, respectively (Table 2). The GGA UUUU motif was absolutely conserved in the overlap region of ORF1a and -1b in HanaV, MénoV, and NséV, further supporting the previously predicted function of this “slippery” sequence in the ribosomal frameshift mechanism required to express the ORF1b-encoded part of pp1ab.

The putative TMDs predicted for the ORF2a-, -3a-, -3b-encoded proteins are conserved among all mesoniviruses, supporting their predicted membrane protein functions (Fig. 2) (2). Downstream of ORF3a/3b, several small ORFs (ranging in size between 120 and 282 nt) were identified. Both the numbers and positions of these small ORFs varied among the four mesoniviruses (Fig. 2). Similarities to other viral and cellular proteins could not be detected for the potential translation products encoded by these ORFs. Also, specific mRNAs from which the ORFs could be expressed were not detected in infected cells (see below), arguing against a functional significance for them.

Phylogenetic analyses of the coding sequences of conserved nonstructural replicase gene-encoded domains, such as RdRp, 3CL^{pro}, Z, and Hel, as well as the putative N and S proteins,

FIG 1 Growth and morphology of mesoniviruses. (A) C6/36 cells were infected with CavV, HanaV, NséV, and MénoV, respectively, at an MOI of 0.001 TCID₅₀ per cell, and the numbers of genome copies per milliliter of cell culture supernatant were measured by specific RT-PCR at the indicated time points postinfection. (B) Negative-staining electron micrograph of NséV (isolate NséV/F7/CI/2004) sedimented by ultracentrifugation through a 36% sucrose cushion.

FIG 2 Mesonivirus genome organization. Shown is a schematic view of CavV, HanaV, MénoV, and NséV genomes. Open reading frames are shown by boxes, with nucleotide positions indicated. Conserved functional domains encoded by the replicase gene (ORF1a and -1b) are indicated by gray boxes. ORF2a is shown in light blue, ORF2b in green, and ORF3a and -3b in yellow. Hydrophobic (putative transmembrane) regions are marked in blue, and predicted signal peptides are shown in red. The light-gray boxes symbolize several small, nonconserved ORFs in the 3'-terminal regions of the four mesonivirus genomes. The solid gray box symbolizes ORF4, while other small ORFs are indicated by open gray boxes.

showed that the newly identified viruses clustered with CavV and NDiV. The data unambiguously identify the four viruses as members of the recently established family *Mesoniviridae* (Fig. 3). While HanaV grouped together with CavV and NDiV, NséV and MoumoV may be members of a sister clade. MénoV was consistently found in a basal relationship to all other mesoniviruses in all our analyses.

Mesoniviruses use two distinct TRS motifs for sgRNA synthesis. We next investigated sg mRNA expression by Northern blotting. Using probes specific for the 5' and 3' genome termini, HanaV, MénoV, and NséV were confirmed to produce three 5'- and 3'-coterminal mRNAs of approximately 20, 4.7, and 1.8 kb, indicating a discontinuous transcription mechanism (Fig. 4 and data not shown). In CavV, another (less abundant) RNA species of about 2.7 kb is produced (2) (Fig. 4).

To obtain more insight into mesonivirus sgRNA synthesis, we sought to identify leader-body fusion sites for all the mRNAs detected by Northern blotting for all four mesoniviruses (Fig. 4). Using RACE-PCR, two different conserved TRS elements were identified. The mRNA that encodes the 2a (S) and 2b (N) proteins in its 5'-terminal (unique) region (tentatively termed sgRNA2), was found to be fused to the 5' leader sequence at ₁₂₄AGACACT CTCCCA, whereas the mRNA that encodes the 3a and 3b proteins in its 5'-terminal region (tentatively termed sgRNA3) used ₂₇TA ATACTACTACTA as a TRS (the nucleotide positions correspond to CavV) (Fig. 5). For the 2.7-kb mRNA detected in CavV-infected cells, a leader-to-body fusion could not be confirmed. Despite

multiple attempts, we failed to obtain evidence for the synthesis of an sgRNA from which ORF4 could be expressed.

Mesonivirus structural protein expression. Investigations of major mesonivirus structural proteins revealed that CavV, HanaV, MénoV, and NséV particles showed similar protein patterns, with eight major proteins of ~77, 57, 25, 23, 20, 19, 18, and 17 kDa being readily detectable (Fig. 6A). To determine the identities of these proteins, Edman sequencing and MALDI-TOF mass spectroscopy were performed using CavV as a representative mesonivirus. N-terminal sequencing revealed that three proteins (p77, p57, and p23) were expressed from ORF2a, suggesting that the putative spike protein is subject to proteolytic cleavage. The proteins were tentatively named S (p77), S1 (p23), and S2 (p57) (Fig. 6B; see Fig. S2 in the supplemental material). S and S1 share the N-terminal sequence ₂₂₄STRIDL, confirming the use of the previously predicted internal signalase cleavage (I) at the ₂₂₀NAHCISTRID₂₂₈ position (2). The N-terminal sequence ₄₃₁WDSSYV was found for the S2 processing product, indicating cleavage at a multibasic cleavage site (₄₂₆KRNKRIWDSS₄₃₄). The site is conserved among mesoniviruses (data not shown) and similar to influenza A virus hemagglutinin (HA) cleavage sites, which are known to be cleaved by furin-like cellular proteases (reviewed in reference 19). Surprisingly, none of the proteins identified in purified virus preparations mapped to residues ₁MINSI-NAHC₂₂₃ of the predicted full-length ORF2a translation product.

The 25-kDa protein could be identified as the putative N protein encoded by ORF2b (Fig. 6B). The four remaining proteins of

TABLE 2 Pairwise identities of replicase domains and complete ORFs of CavV, NDiV, HanaV, NsÉV, MénoV, and MoumoV

Gene	Virus	Size [nt (aa)]				
		NDiV	HanaV	MénoV	NsÉV	MoumoV
3CL ^{pro}	CavV	90.5 (96.4)	83.2 (90.5)	67.0 (72.7)	72.0 (79.3)	
	NDiV		83.0 (89.5)	66.0 (72.4)	72.5 (78.6)	
	HanaV			69.8 (71.4)	74.3 (79.3)	
	MénoV				68.3 (72.4)	
RdRp	CavV	92.9 (98.5)	83.4 (86.8)	72.6 (76.6)	76.0 (80.7)	78.1 (83.6)
	NDiV		82.7 (86.6)	72.3 (76.4)	75.3 (80.7)	77.6 (83.6)
	HanaV			71.1 (72.6)	75.4 (79.4)	76.3 (81.0)
	MénoV				70.8 (75.1)	73.4 (78.3)
Z	CavV	94.8 (95.5)	87.3 (86.5)	77.0 (77.8)	81.3 (82.0)	76.4 (82.0)
	NDiV		89.1 (88.8)	76.7 (77.8)	80.9 (80.9)	77.9 (83.1)
	HanaV			75.2 (74.4)	79.4 (79.8)	78.3 (80.9)
	MénoV				76.3 (76.7)	76.3 (74.4)
Hel	CavV	92.2 (98.8)	84.9 (90.4)	78.1 (84.9)	74.8 (81.4)	75.2 (82.0)
	NDiV		86.2 (90.7)	77.7 (84.9)	74.8 (82.6)	76.2 (82.3)
	HanaV			74.9 (79.9)	74.6 (79.9)	75.8 (80.8)
	MénoV				73.4 (81.7)	75.0 (81.7)
ExoN	CavV	92.3 (94.2)	84.3 (86.0)	67.4 (64.3)	73.4 (70.2)	72.4 (67.4)
	NDiV		85.3 (86.8)	68.1 (65.1)	74.2 (69.4)	73.8 (67.4)
	HanaV			68.0 (64.7)	74.2 (70.9)	74.2 (66.7)
	MénoV				69.6 (67.1)	70.0 (65.5)
NMT	CavV	92.0 (94.0)	87.0 (87.6)	69.6 (67.7)	69.4 (67.8)	69.2 (68.4)
	NDiV		85.8 (89.3)	70.9 (68.5)	69.5 (69.1)	69.3 (70.5)
	HanaV			71.9 (69.4)	69.8 (67.4)	70.0 (70.5)
	MénoV				67.7 (61.2)	68.2 (63.4)
OMT	CavV	91.4 (94.6)	80.4 (76.9)	69.6 (66.1)	75.1 (69.0)	73.8 (70.7)
	NDiV		80.2 (76.4)	70.4 (67.8)	75.5 (70.7)	74.5 (70.7)
	HanaV			67.2 (63.8)	70.5 (65.3)	70.1 (63.2)
	MénoV				70.5 (65.8)	68.2 (64.2)
ORF1a	CavV	88.4 (90.0)	79.5 (80.4)	58.5 (49.1)	64.3 (59.0)	
	NDiV		80.0 (81.6)	58.5 (48.7)	64.2 (59.1)	
	HanaV			58.8 (49.1)	64.8 (59.0)	
	MénoV				59.5 (50.3)	
ORF1b	CavV	92.6 (96.0)	83.7 (85.3)	71.0 (70.4)	74.3 (74.2)	
	NDiV		83.9 (85.7)	71.2 (70.8)	74.4 (74.9)	
	HanaV			69.8 (68.3)	73.6 (70.0)	
	MénoV				69.7 (68.7)	
ORF2a	CavV	90.7 (87.7)	84.1 (80.3)	66.5 (57.1)	68.8 (61.2)	
	NDiV		85.1 (81.7)	66.7 (57.8)	70.3 (62.7)	
	HanaV			67.1 (58.2)	70.0 (63.7)	
	MénoV				64.5 (56.9)	
ORF2b	CavV	88.8 (90.2)	78.3 (78.6)	60.2 (50.9)	65.8 (58.3)	
	NDiV		79.3 (78.4)	60.3 (50.7)	66.9 (58.8)	
	HanaV			62.9 (55.1)	66.1 (58.3)	
	MénoV				59.4 (52.6)	
ORF3a	CavV	91.2 (93.1)	82.2 (83.6)	63.7 (57.9)	73.4 (68.5)	
	NDiV		83.6 (84.9)	64.2 (58.4)	72.0 (69.1)	
	HanaV			65.4 (58.1)	69.8 (69.1)	
	MénoV				62.9 (56.5)	

TABLE 2 (Continued)

Gene	Virus	Size [nt (aa)]				
		NDiV	HanaV	MénoV	NsÉV	MoumoV
ORF3b	CavV	93.7 (90.6)	89.2 (81.2)	63.9 (47.5)	71.3 (60.7)	
	NDiV		86.0 (78.6)	64.9 (50.0)	68.9 (58.2)	
	HanaV			64.2 (49.2)	70.5 (58.2)	
	MénoV				61.5 (50.8)	

20, 19, 18, and 17 kDa were mapped to ORF3a, encoding the putative M protein (Fig. 6B), which was previously predicted to contain a signal peptidase cleavage site and membrane-spanning domains (2). N-terminal sequence data obtained for all four ORF3a-derived protein species consistently revealed cleavage at the predicted site, ₁₄VAMSI₂₁AETD₂₁, suggesting that the slightly different molecular masses observed for these proteins resulted from differential glycosylation patterns. This hypothesis was further supported by predictions of *N*-glycosylation sites for the putative M and S proteins (Fig. 6B). To study *N*-linked glycosylation of mesonivirus structural proteins, purified virion proteins of CavV, HanaV, MénoV, and NsÉV were treated with peptide-*N*-glycosidase F (PNGase F) and analyzed by SDS-PAGE. The data provided evidence that the S, S2, and M proteins were posttranslationally modified by *N*-linked glycosylation (Fig. 6A and B), while there was no evidence for glycosylation of the 2b protein, consistent with its presumed role as the mesonivirus N protein. No proteins corresponding to ORF3b, ORF4, or other small ORFs were detected.

To test for serological cross-reactivity between CavV, HanaV, MénoV, and NsÉV, a Western blot analysis with rabbit anti-CavV/NsÉV serum was performed. The S and M proteins of all viruses were equally detected by the antiserum, suggesting conservation of major antigenic sites (Fig. 7A). Additionally, anti-peptide antibodies raised against specific epitopes (derived from coding sequences of CavV structural proteins) of the predicted ORF2a-, ORF2b-, and ORF3a-encoded proteins were generated (rabbit anti-CavV S, N, and M sera) (Table 3). Furthermore, epitope-specific antibodies directed against sequences present in the N-terminal 223 amino acids (aa) of ORF2a (rabbit anti-CavV p2aX serum) and against the potential gene products encoded by ORF3b and ORF4 (rabbit anti-CavV p3b and p4 sera) were generated to investigate if these ORFs are functional. As shown in Fig. 7, the respective epitope-specific polyclonal antibodies efficiently detected the CavV N, S, S1, S2, and M proteins (Fig. 7B to D). Deglycosylation of proteins resulted in stronger reactivity of rabbit anti-CavV S serum, especially toward the S2 subunit, providing evidence for *N*-linked glycosylation at one of the predicted sites (₅₂₆I₅₂₆ENTTNTL₅₃₃; boldface indicates the predicted *N*-glycosylated residue) (Fig. 7C), which overlaps one of the epitopes used to raise CavV S-specific antisera, ₅₂₄NHIENTTNTLENRIN₅₃₉ (Table 3 and Fig. 6B). The S1 subunits of the various mesoniviruses were not equally well detected by the CavV S-specific antiserum, probably reflecting the poor conservation among mesoniviruses of the N-terminal part of epitope 1, especially in the case of MénoV (Table 3 and Fig. 6B). Few differences between glycosylated and deglycosylated proteins were observed with rabbit anti-CavV M serum, indicating that, in this case, the epitopes used to raise rabbit antisera were not masked by *N*-linked glycosylation (Fig. 7D). The CavV M protein-specific antiserum was found to

FIG 3 Phylogenetic relationships of mesoniviruses. Shown is a maximum-likelihood phylogenetic tree of mesoniviruses and representative nidoviruses based on amino acid sequences of the conserved pp1ab domains RdRp, 3CL^{pro}, Z, Hel, and the putative N and S proteins. Analyses were performed using the BLOSUM62 substitution model without optimization and with an estimated proportion of invariable sites and an estimated gamma distribution parameter with 1,000 bootstrap replicates.

cross-react with the HanaV M protein but failed to detect the M protein homologs from MénoV and NséV, which is in agreement with the poor conservation of the two CavV M protein epitopes used to raise this particular antiserum (Table 3). In contrast, the observed cross-reactivity of the antiserum for the CavV and HanaV M proteins can be explained by the complete conservation of epitope 2 in both viruses. The rabbit anti-CavV N serum was able to detect the putative N proteins from all four mesoniviruses, most likely because epitope 2 was very well conserved among these viruses. Experiments using rabbit anti-CavV p2aX, p3b, and p4 failed to provide evidence for expression of the first 223 aa of

ORF2a or the predicted proteins p3b and p4 (data not shown). Overall, the serological (cross-reactivity) data are in good agreement with the observed phylogenetic relationships between the four mesoniviruses and confirm the identities of mesonivirus structural proteins as revealed by N-terminal sequence analysis and mass spectroscopy.

DISCUSSION

In this study, we identified four novel viruses in mosquitoes originating from the same area as CavV. The four viruses were classified as belonging to the newly established family *Mesoniviridae*,

FIG 4 Mesonivirus genomic and subgenomic RNA synthesis. Northern blot analysis of viral RNA isolated from infected C6/36 cells. Digoxigenin (DIG)-labeled probes specific for the 3' ends of CavV, HanaV, MénoV, and NséV, respectively, were used for detection. RNA from noninfected C6/36 cells was used as a control. A DIG-labeled RNA was used as a size marker (M), with sizes given in nucleotides on the right and sizes of mRNAs given in kilobases on the left. For the detection of viral genomic RNA, a longer exposure time was used (shown at the top). The positions of genome RNA and the two major sgRNAs (mRNA2 and mRNA3) are indicated on the left.

and important biological properties, such as host range, growth characteristics, genome organization and expression, and phylogenetic relationships, as well as viral RNAs and structural proteins, were analyzed.

The mosquitoes used in our study were sampled along an anthropogenic disturbance gradient ranging from a primary rainforest to human settlements, wherein CavV was detected at a prevalence of 9.3% and was shown to have originated in the primary rainforest, from where it spread into human settlements (1, 2). This spread was found to be associated with a significant increase in CavV prevalence (from 4.1 to 26.5%) and a concomitant decrease in genetic diversity (from 1.4 to 0.1%) (2). In contrast, NséV, HanaV, MénoV, and MoumoV were consistently less prevalent, ranging from 0.5 to 2.8% along the gradient. This low prevalence precluded further studies of prevalence and diversity patterns along the gradient. Given the large number of insect species included in the study, we consider it unlikely that a (potentially existing) pronounced host specificity of HanaV, MénoV, NséV, and MoumoV was the cause for the observed low prevalence of these four viruses in our samples. As CavV was detected much more frequently in these samples, future studies might explore the possibility of CavV competing with these novel, or other, mesoniviruses in insects. The presence of other viruses in the samples may also have influenced mesonivirus prevalence (1, 12, 20, 21).

Differences in viral detection rates could also be explained by differential intrinsic capabilities of the viruses to replicate. Virus growth characteristics *in vitro*, including replication kinetics and peak concentrations, proved to be very similar for all the viruses studied. Interestingly, however, CavV induced a strong CPE, while all other mesoniviruses were significantly less cytopathogenic, causing only minimal morphological changes in C6/36 cells. Even though cell cultures provide an imperfect surrogate of the virus life cycle in insects, the data may indicate potential differences in pathogenicity, and thus transmission or maintenance,

FIG 5 Mesonivirus TRSs. Shown is an overview of TRS core sequences and leader-body fusion sites determined in this study for mRNA2 and -3 of CavV, HanaV, MénoV, and NséV, respectively. The actual fusion sites are shown in blue and the conserved TRS core sequences in red. Production of the two sgRNAs detected in mesonivirus-infected cells involves (i) two different transcription-regulating sequences (TRS1 and -2) and (ii) two different fusion sites of leader and body sequences. As a result, the two major sgRNAs contain 5' leader sequences of different lengths (~120 versus ~20 nucleotides). The genomic positions of leader and body TRSs are indicated.

in insects (22–25). In a first set of experiments using a wide range of vertebrate cell lines, we obtained no evidence for replication of any mesonivirus in any of the cell lines tested. The presumably multiple requirements for mesonivirus replication in specific cell types remain to be addressed in future studies. However, our data suggest that one of these requirements involves cleavage by specific cellular proteases at a multibasic cleavage site conserved in the mesonivirus S protein. Cleavage at these sites likely activates the receptor-binding and fusion activities of the S protein, which in turn promotes efficient virus entry into specific target cells.

Our studies of mesonivirus morphology showed that virions of the newly identified mesoniviruses were similar to those observed for CavV (2). Some of the HanaV and MénoV virions appeared to contain a lower number of spike projections. It remains to be determined if this difference reflects an intrinsic instability of (some) mesonivirus spike proteins or less efficient integration of the protein in the viral envelope, resulting in the different morphologies observed for different mesoniviruses (2, 3) (Fig. 1B to D) and possibly providing an explanation for the apparent absence of surface projections in preparations of NDIV (3).

Full genome comparisons of four distinct mesoniviruses revealed conserved properties, such as genome sizes of around 20 kb; conservation of ORF1a/b, -2a/b, and -3a/b; and the presence of polypeptide domains in the order TM-3CL^{PTO}-TM-RFS-RdRp-Z-Hel-ExoN-NMT-OMT (26). In contrast to other nidoviruses,

FIG 6 Mesonivirus structural proteins. (A) SDS-PAGE analysis of the protein content of virus particles purified from cell culture supernatants of infected cells by gradient ultracentrifugation. The proteins were analyzed either directly or after deglycosylation using PNGase F (– and +, respectively). Proteins were stained with Coomassie blue R-250. (B) The identities of CavV virion proteins were analyzed by Edman degradation and MALDI-TOF mass spectroscopy. Sequences identified by Edman degradation are shown in red, and MALDI-TOF sequence coverage is shown in yellow. Specific ORFs shown to encode the respective proteins are given on the right, together with the observed molecular masses of the proteins. Predicted N-linked glycosylation sites in the S and M proteins are indicated by brown triangles. The approximate positions of epitopes recognized by the antibodies used in this study are illustrated by green boxes.

species classification criteria have not been established for mesoniviruses. We therefore compared genetic distances of conserved polyprotein domains that have recently been approved as species classification criteria for the related CoV (26). Intergeneric amino acid identities between the RdRp genes of *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus*, and *Deltacoronavirus* ranged between 48.0 and 62.3%. Pairwise identities of HanV, MénoV, NséV, CavV, and NDiV ranged between 86.9 and 75.6% and thus were within the intrageneric, but not the intergeneric, distance range as defined for CoV. Additionally, our data show that the mesoniviruses characterized in this study are serologically related. Among CoV, significant cross-reactivity is generally limited to viruses belonging to the same genus (26, 27). By analogy, our serological data suggest that the mesoniviruses under study all belong to the same genus, *Alphamesonivirus*.

The current demarcation of CoV species has essentially been based on the presence of more than 90% amino acid sequence identity in the seven most conserved polyprotein pp1ab domains, i.e., ADP-ribose-1"-phosphatase (ADRP), 3CL^{pro}, RdRp, Hel, ExoN-NMT, NendoU, and OMT. Five of these domains (the exceptions are ADRP and NendoU) are conserved in mesoniviruses.

Pairwise identities between these pp1ab domains of alphamesonivirus 1 (CavV and NDiV), HanaV, MénoV, and NséV never exceeded 82.6%, suggesting that, if the same criteria approved for the *Coronavirinae* are applied to members of the *Mesoniviridae*, four different mesonivirus species have now been described, three of which were newly identified in the present study. The amino acid identity of MoumoV, from which only a 7.9-kb fragment covering four of the conserved pp1ab domains could be retrieved, was up to 87.3% with NséV (its closest relative). Because the pairwise identities between established species of the genera *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* (two approved species only), and *Deltacoronavirus* never exceeded 90.2%, 90.2%, 70.9%, and 92.5%, respectively, MoumoV may or may not represent an additional, fifth tentative species, pending more extensive sequence information.

Extensive phylogenetic analyses of mesonivirus relationships suggested that MénoV may represent a sister taxon to all other mesoniviruses. MénoV had the shortest genome of all mesoniviruses and lacked ORF4, a gene conserved in all other mesoniviruses. Gene loss along this long lineage appears to be more likely than gene acquisition along the shorter branch connecting the common mesonivirus ancestor with the ancestor of the ORF4-containing sister taxon to MénoV. Gain and loss of genes in the structural and accessory gene region is regularly observed in CoVs and ToVs (28–30) (for a recent review, see reference 31). The presence of an ExoN domain in mesoniviruses may indicate an intrinsic potential of this nidovirus lineage for genome expansion that also may have facilitated the acquisition of additional protein genes (3). Sequence analyses of yet-to-be-discovered mesoniviruses, potentially also from other genera, is expected to provide further insight into the evolution of this new virus family.

On the same account, we should mention that the topology of the nidovirus tree found here and earlier (2) deviated from that of Nga and colleagues (3) in the basal topology of the clade of viruses with genomes of >20 kb. Our phylogenetic analyses of single gene data sets placed RoV as a solitary clade on one side of a basal bifurcation and all other medium-size and large nidoviruses on the other side. Because of the huge phylogenetic distance between nidovirus taxa, it is difficult to infer fully resolved phylogenies for basal nodes. We therefore focused on highly conserved protein domains, thus giving up information on less conserved genome regions, which is required to achieve maximum apical resolution and was used previously for phylogenetic analyses of the *Coronavirinae* (32). Without proposing that the deep node topology determined here is the only correct one, a number of genome features not included in the inference of phylogenies support our proposal. First, RoVs have the smallest number of ORFs downstream of the replicase polyprotein (RoVs have two [33], mesoniviruses at least three, ToVs three or four [34], and CoVs up to 12 [7]), indicating a potential evolutionary trend that would be in line with our topology. Second, unlike other nidovirus sgRNAs, RoV sgRNAs do not contain leader sequences (35), providing another example of a special phylogenetic position of this nidovirus lineage. Third, the RoV glycoprotein contains six transmembrane domains in contrast to just one in all other nidoviruses of >20 kb, again placing RoV on a separate branch. Fourth, the N protein gene is located upstream of the glycoprotein ORF in RoV, while in mesoniviruses it overlaps the S gene ORF and in CoV and ToV it is located downstream of the glycoprotein (S) gene ORF (see Fig. S2

FIG 7 Western blot analysis of mesonivirus proteins. Shown is Western blot analysis of viral proteins present in CavV, HanaV, MénoV, and NsêV virions purified by gradient ultracentrifugation. The proteins were analyzed either directly (–) or after deglycosylation using PNGase F (+) and detected using polyclonal rabbit anti-CavV/NsêV serum (A), rabbit anti-CavV N serum (B), rabbit anti-CavV S serum (C), and rabbit anti-CavV M serum (D). In panels A and C, proteins shown in the upper blot were separated by SDS-PAGE on a 12% polyacrylamide gel and those in the lower blot on a 14% polyacrylamide gel.

in the supplemental material). This gradual shift of gene positions might indicate another trend in support of our phylogeny. Finally, RoVs are associated with the evolutionarily oldest hosts (*Crustacea*) (36), mesoniviruses with insects, bafiniviruses and toroviruses with fish and vertebrates, and CoVs with vertebrates only; the most basal CoV taxon is associated with birds and the more apical taxa with mammals. This may suggest cosegregation of major evolutionary nidovirus lineages with major evolutionary lineages in the animal kingdom. The characterization of yet-to-be-discovered nidoviruses in other hosts should help to test or validate these speculations.

The use of two different TRS elements in mesoniviruses adds to the variability of nidovirus sgRNA synthesis and has not been described previously for any other nidovirus. The most prominent mesonivirus protein was shown to be the M protein, expressed from sgRNA3, which is produced in larger amounts than sgRNA2. It remains to be determined whether the observed gradient for viral RNA synthesis (sgRNA3>sgRNA2>RNA1) is solely caused by the (postulated) discontinuous extension process of subgenomic minus strands (as demonstrated previously for CoVs and ArVs [10, 37]) or, in addition, by specific features of TRS2 more efficiently promoting the translocation of the replication-transcription complex to the upstream leader TRS position during negative-strand synthesis. The lower expression levels of proteins expressed from sgRNA2 (S and N) correlates with the smaller amounts of mRNA2 produced in virus-infected cells. Additionally, the leaky scanning mechanism predicted to be required to initiate translation of the viral N protein at a downstream AUG codon can be expected to further reduce N protein expression. Interestingly, and similar to what was reported previously for the bafinivirus white bream virus (WBV) (11) and several other nidovirus sgRNAs (38), the actual fusion sites identified in mesonivirus sgRNAs were found to be located a few nucleotides upstream of

the leader TRS, suggesting that, following the TRS-guided template switch during minus-strand sgRNA synthesis, RdRp is able to extend the nascent RNA strand from a (partially) unpaired 3'-terminal end. The positions of the fusion sites identified here differed slightly from those identified in our initial study (2). Three potential leader-body fusion sites were previously identified by RT-PCR for CavV RNA2 (4.7 kb), RNA3 (2.7 kb), and RNA4 (1.8 kb). The fusion sites of CavV RNA2 and -3 are in agreement with data from the recent study and correspond to TRS1 (previously reported leader-body fusion sites were slightly displaced [2]). However, gene RACE PCR could not confirm the fusion site proposed earlier for CavV RNA4. Further studies are required to investigate if this aberrant fusion was an RT-PCR artifact or indicates a minor sgRNA species. TRS2 was not identified previously. Our new data show that the previously used RT-PCR sense primer binds to a sequence downstream of the TRS2 leader-body fusion site, thus providing a plausible explanation for the discrepancy between the two studies.

Our data suggest that mesonivirus virions contain eight major structural protein species, i.e., a nonglycosylated p25 (a putative N protein), four forms of a glycosylated membrane protein (a putative M protein, p17, p18, p19, and p20), and three ORF2a-encoded glycoproteins (a putative S protein and two S protein subunits), that are expressed as a precursor, tentatively named S protein (p77), which is (partially) cleaved into two subunits of 23 kDa and 57 kDa, tentatively named S1 and S2. A cleavage product corresponding to the N-terminal 223 amino acid residues of the ORF2a-encoded protein could not be detected in virions, indicating no (or inefficient) incorporation of this N-terminal domain into virus particles. NDiV virion proteins were found to contain an S protein (p2a; 77 kDa) with the N-terminal sequence STRID, but no cleavage products were detected (3). For p2a, a relatively weak protein band was detected by SDS-PAGE analysis of NDiV

TABLE 3 Comparison of epitope antibody sequences of CavV, HanaV, MénoV, and NséV

Antibody	ORF	Virus	Sequence of epitope ^a :	
			1	2
Anti-CavV-S	ORF2a	CavV	EKALKRNKRWDSSSYVC	NHIENTTNTLENRINV
		HanaV	. . . Q . . H	SE . K S . . AI
		MénoV	KAENR . Q	EQVR VQTSQ . NI
		NséV	TQSKI . KR	EQVK S . . S . SI
Anti-CavV-N	ORF2b	CavV	QPGTSKQNRPTKNSK	ARNHMGWRKNEKTGS
		HanaV	K . . P . R . . QQP . -TT	 R
		MénoV	SS-QL . . . KQSNPRV	 R . . . N . .
		NséV	T . KQG . . . -N . N . GT	 R
Anti-CavV-M	ORF3a	CavV	SAETDPEVVSPPSKLC	RQSDGSYTLLP-GRSYR
		HanaV	A . DS . VAPS . . A . . .	 -
		MénoV	. . LA . DAPTNTPT . .	. DNH . . IQ . . . I . K .
		NséV	AQ . SSNGGAN-ST . . .	RK . N II . -PTIR
Anti-CavV-p3b	ORF3b	CavV	CIRT-QHQRQNNCTTSP	I-RSSQDDRID-KLQSR
		HanaV	 - . P . S . . H . I . .	. - G . . A
		MénoV	. MNQAPPKESVPSR . . .	SNYYP . E - - - N . IDG
		NséV	 AH . TS . TT . I . . .	. RLSPQQS . LEM . NQ . LN .
Anti-CavV-p4	ORF4	CavV	KRSAHVYVSGLGPNNR	FLYIR-TQGLEQADHSC
		HanaV	. . I	 -
		MénoV	/	/
		NséV	. Y	 YARP . TSRTFKLN
Anti-CavV-p2aX	ORF2a	CavV	ALRNKTGPPKILKPE	LNNYHHQERQNKAT
		HanaV	. PHV . ISSQ . P . T . .	
		MénoV	NKSVNLST . FLAI -	- - - TNPLL - - - - - S
		NséV	QTNPQCHSDA . . I - .	- - - AKPEL - - - - - V

^a Dashes, ORF not present; periods, identical aa; slashes, gap.

virion proteins, and the apparent absence of S1 and S2 may have been caused by too low protein concentrations in the preparations analyzed. Cleavage of the S protein into two subunits is in agreement with findings in several other nidoviruses (39–42). Interestingly, similar locations of cleavage sites and absence of the most N-terminal cleavage product were also reported for the envelope glycoproteins of RoV (40). Thus, purified virions of yellow head virus (YHV) were shown to contain two envelope glycoproteins of 116 kDa (gp116) and 64 kDa (gp64), both proteolytically cleaved from an ORF3-encoded protein precursor of 185.7 kDa. Interestingly, and similar to our findings, the predicted N-terminal cleavage product of 227 aa (25.4 kDa) has not been identified for YHV. The N-terminal cleavage site of gp116 in YHV and the related gill-associated virus (GAV) were reported to be ₂₂₇PAFAI/TILS₂₃₄ (40) and ₂₂₅PTFAI/KEPE₂₃₂, respectively.

Interestingly, the N-terminal signal peptidase cleavage site of the mesonivirus S1 subunit (₂₂₀NAHC|STRID₂₂₈) was identified as also being conserved between YHV and GAV (₅₀₉EIYA|STRA/HD₅₁₃) (the genome positions refer to GAV, and residues conserved among RoV gp116 and mesonivirus S proteins are underlined). The cleavage site may have become functional in insects after the N ORF had been moved from its position upstream of the structural glycoprotein gene (ORF3) in RoV to a position further downstream in the genome, now overlapping the S ORF (ORF2a) in mesoniviruses. This process would have involved a reduction of genome size in mesoniviruses while maintaining specific mechanisms suitable to regulate the expression levels of these structural proteins.

CoV S proteins are known to form homotrimers, with their S1 subunits forming the globular membrane-distal part containing the receptor-binding domain and S2 subunits forming the mem-

brane-anchored stalk essential for fusion activity (43–45). This conformation is also suggested for the S proteins of ToV, and it seems reasonable to predict a similar structure for the mesonivirus S protein, as the C-terminal part of the S2 subunit is predicted to contain a hydrophobic region, including two putative heptad repeat (HR) regions (HR1 [₄₉₉N-N₅₃₅] and HR2 [₈₃₇E-A₈₅₂], using COILS/PCOILS [<http://toolkit.tuebingen.mpg.de/pcoils>]), suggesting that mesonivirus S protein subunits oligomerize and that this process involves interactions between amphipathic α -helices.

Although ORF3b and ORF4 were found to be conserved in mesoniviruses, no proteins expressed from the two ORFs could be detected by SDS-PAGE or Western blotting. ORF3b was predicted previously to encode a protein with membrane-spanning domains (2), suggesting a membrane protein function. Given the downstream position of ORF3b on mRNA3, the predicted protein might be expressed at very low levels (below the detection limit of our assays). The putative protein encoded by ORF4 in CavV, HanaV, and NséV, for which not even an abundantly expressed sgRNA could be detected in our experiments, may also have a very low expression level. In GAV, a 9.6-kDa protein was reported to be expressed at a low level from ORF4 in GAV-infected shrimp tissue (46). Accordingly, further studies may help clarify if these downstream ORFs are expressed in mosquitoes infected with specific mesoniviruses.

In summary, we discovered a high level of mesonivirus diversity in a genetically rather narrow sample of mosquitoes from only one geographic region. Insects from other taxa and geographic regions should be tested for mesoniviruses to help answer important questions related to the molecular biology, phylogeny, host range, and ecological factors that influence the abundance, prevalence, and spread of this fascinating new virus family.

ACKNOWLEDGMENTS

We thank Günter Lochnit, Justus Liebig University of Giessen, and Marc Sylvester, University of Bonn, for protein sequence data. For technical assistance, we thank Pascal Trippner. We thank the Ministry of the Environment and Forests and the Ministry of Research, the directorship of the Taï National Park, the Swiss Research Center in Côte d'Ivoire, and the Max Planck Society for long-term support.

The project was supported by the BonFor program of the University of Bonn (grant agreement no. O-156-0006) and ANTIGONE (grant agreement no. 278976).

REFERENCES

- Junglen S, Kurth A, Kuehl H, Quan PL, Ellerbrok H, Pauli G, Nitsche A, Nunn C, Rich SM, Lipkin WI, Briese T, Leendertz FH. 2009. Examining landscape factors influencing relative distribution of mosquito genera and frequency of virus infection. *Ecohealth* 6:239–249.
- Zirkel F, Kurth A, Quan PL, Briese T, Ellerbrok H, Pauli G, Leendertz FH, Lipkin WI, Ziebuhr J, Drosten C, Junglen S. 2011. An insect nidovirus emerging from a primary tropical rainforest. *mBio* 2:e00077–00011. doi:10.1128/mBio.00077-11
- Nga PT, del Carmen Parquet M, Lauber C, Parida M, Nabeshima T, Yu F, Thuy NT, Inoue S, Ito T, Okamoto K, Ichinose A, Snijder EJ, Morita K, Gorbalenya AE. 2011. Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. *PLoS Pathog.* 7:e1002215. doi:10.1371/journal.ppat.1002215
- Lauber C, Ziebuhr J, Junglen S, Drosten C, Zirkel F, Nga PT, Morita K, Snijder EJ, Gorbalenya AE. 2012. Mesoniviridae: a proposed new family in the order Nidovirales formed by a single species of mosquito-borne viruses. *Arch. Virol.* 157:1623–1628.
- Ziebuhr J. 2006. The coronavirus replicase: insights into a sophisticated enzyme machinery. *Adv. Exp. Med. Biol.* 581:3–11.
- Firth AE, Brierley I. 2012. Non-canonical translation in RNA viruses. *J. Gen. Virol.* 93:1385–1409.
- Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* 117:17–37.
- Ivanov KA, Hertzog T, Rozanov M, Bayer S, Thiel V, Gorbalenya AE, Ziebuhr J. 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 101:12694–12699.
- Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* 331:991–1004.
- Pasternak AO, Spaan WJ, Snijder EJ. 2006. Nidovirus transcription: how to make sense? *J. Gen. Virol.* 87:1403–1421.
- Schutze H, Ulferts R, Schelle B, Bayer S, Granzow H, Hoffmann B, Mettenleiter TC, Ziebuhr J. 2006. Characterization of White bream virus reveals a novel genetic cluster of nidoviruses. *J. Virol.* 80:11598–11609.
- Junglen S, Kopp A, Kurth A, Pauli G, Ellerbrok H, Leendertz FH. 2009. A new flavivirus and a new vector: characterization of a novel flavivirus isolated from uranotaenia mosquitoes from a tropical rain forest. *J. Virol.* 83:4462–4468.
- Biesold SE, Ritz D, Gloza-Rausch F, Wollny R, Drexler JF, Corman VM, Kalko EK, Oppong S, Drosten C, Muller MA. 2011. Type I interferon reaction to viral infection in interferon-competent, immortalized cell lines from the African fruit bat *Eidolon helvum*. *PLoS One* 6:e28131. doi:10.1371/journal.pone.0028131
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649.
- Katoh K, Standley DM. 16 January 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* [Epub ahead of print.]. doi:10.1093/molbev/mst010
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59:307–321.
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685.
- Refaie FM, Esmat AY, Mohamed AF, Mohamed WA. 2004. The effect of chemical inactivation of bovine viral diarrhoea virus with beta-propiolactone and binary ethyleneimine on plasma proteins and coagulation factors. *Egypt J. Immunol.* 11:9–20.
- Palese P, Shaw ML. 2007. *Orthomyxoviridae: the viruses and their replication*, p 1648–1689. In Knipe DM, Howley PM (ed), *Fields virology*, 5th ed, vol 2. Lippincott Williams & Wilkins, Philadelphia, PA.
- Marklewitz M, Handrick S, Grasse W, Kurth A, Lukashev A, Drosten C, Ellerbrok H, Leendertz FH, Pauli G, Junglen S. 2011. Gouleako virus isolated from West African mosquitoes constitutes a proposed novel genus in the family Bunyaviridae. *J. Virol.* 85:9227–9234.
- Quan PL, Junglen S, Tashmukhamedova A, Conlan S, Hutchison SK, Kurth A, Ellerbrok H, Egholm M, Briese T, Leendertz FH, Lipkin WI. 2010. Moussa virus: a new member of the Rhabdoviridae family isolated from *Culex decens* mosquitoes in Cote d'Ivoire. *Virus Res.* 147:17–24.
- Flatz L, Rieger T, Merkler D, Berghaler A, Regen T, Schedensack M, Bestmann L, Verschoor A, Kreutzfeldt M, Bruck W, Hanisch UK, Gunther S, Pinschewer DD. 2010. T cell-dependence of Lassa fever pathogenesis. *PLoS Pathog.* 6:e1000836. doi:10.1371/journal.ppat.1000836
- Kraus AA, Raftery MJ, Giese T, Ulrich R, Zawatzky R, Hippenstiel S, Suttrop N, Kruger DH, Schonrich G. 2004. Differential antiviral response of endothelial cells after infection with pathogenic and nonpathogenic hantaviruses. *J. Virol.* 78:6143–6150.
- Lackner T, Muller A, König M, Thiel HJ, Tautz N. 2005. Persistence of bovine viral diarrhoea virus is determined by a cellular cofactor of a viral autoprotease. *J. Virol.* 79:9746–9755.
- Schneider DS, Ayres JS. 2008. Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. *Nat. Rev. Immunol.* 8:889–895.
- de Groot RJ, Cowley JA, Enjuanes L, Faaborg KS, Perlman S, Rottier PJM, Snijder EJ, Ziebuhr J, Gorbalenya AE. 2012. Order Nidovirales, p 785–834. In King AMQ, Adams MJ, Carstens EB, Lefkowitz FJ (ed), *Virus taxonomy*. Elsevier, Amsterdam, The Netherlands.
- Spaan WJ, Cavanagh D, Horzinek MC. 1990. Coronaviruses, p 359–379. In Regenmortel MHV, Neurath AR. (ed), *Immunochemistry of viruses*, vol 2. Elsevier, Amsterdam, The Netherlands.
- Cornelissen LA, Wierda CM, van der Meer FJ, Herrewegh AA, Horzinek MC, Egberink HF, de Groot RJ. 1997. Hemagglutinin-esterase, a novel structural protein of torovirus. *J. Virol.* 71:5277–5286.
- Lorusso A, Decaro N, Schellen P, Rottier PJ, Buonavoglia C, Haijema BJ, de Groot RJ. 2008. Gain, preservation, and loss of a group 1a coronavirus accessory glycoprotein. *J. Virol.* 82:10312–10317.
- Ziebuhr J. 2008. Coronavirus replicative proteins, p 65–81. In Perlman S, Gallagher T, Snijder EJ (ed), *Nidoviruses*. ASM Press, Washington, DC.
- Narayanan KC, Huang C, Makino S. 2008. Coronavirus accessory proteins. ASM Press, Washington, DC.
- de Groot RJ, Baker SC, Baric R, Enjuanes L, Gorbalenya AE, Holmes KV, Perlman S, Poon L, Rottier PJM, Talbot PJ, Woo PCY, Ziebuhr J. 2012. Family Coronaviridae, p 806–828. In King AMQ, Adams MJ, Carstens EB, Lefkowitz E (ed), *Virus taxonomy*. Elsevier, Amsterdam, The Netherlands.
- Cowley JA, Walker PJ. 2002. The complete genome sequence of gill-associated virus of *Penaeus monodon* prawns indicates a gene organisation unique among nidoviruses. *Arch. Virol.* 147:1977–1987.
- Smits SL, van Vliet AL, Segeren K, el Azzouzi H, van Essen M, de Groot RJ. 2005. Torovirus non-discontinuous transcription: mutational analysis of a subgenomic mRNA promoter. *J. Virol.* 79:8275–8281.
- Cowley JA, Dimmock CM, Walker PJ. 2002. Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. *J. Gen. Virol.* 83:927–935.
- Cowley JA, Dimmock CM, Spann KM, Walker PJ. 2000. Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. *J. Gen. Virol.* 81:1473–1484.
- Sawicki SG, Sawicki DL, Siddell SG. 2007. A contemporary view of coronavirus transcription. *J. Virol.* 81:20–29.
- den Boon JA, Kleijnen MF, Spaan WJ, Snijder EJ. 1996. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J. Virol.* 70:4291–4298.
- Cavanagh D. 1995. The coronavirus surface glycoprotein, p 73–113. In Siddell SG (ed), *The Coronaviridae*. Plenum Press, New York, NY.

40. Jitrapakdee S, Unajak S, Sittidilokratna N, Hodgson RA, Cowley JA, Walker PJ, Panyim S, Boonsaeng V. 2003. Identification and analysis of gp116 and gp64 structural glycoproteins of yellow head nidovirus of *Penaeus monodon* shrimp. *J. Gen. Virol.* **84**:863–873.
41. Snijder EJ, Horzinek MC. 1995. The molecular biology of toroviruses, p 219–238. *In* Siddell SG (ed), *The Coronaviridae*. Plenum Press, New York, NY.
42. Sturman LS, Ricard CS, Holmes KV. 1985. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J. Virol.* **56**:904–911.
43. Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* **77**:8801–8811.
44. de Groot RJ, Luytjes W, Horzinek MC, van der Zeijst BA, Spaan WJ, Lenstra JA. 1987. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. *J. Mol. Biol.* **196**:963–966.
45. Suzuki H, Taguchi F. 1996. Analysis of the receptor-binding site of murine coronavirus spike protein. *J. Virol.* **70**:2632–2636.
46. Dhar AK, Cowley JA, Hasson KW, Walker PJ. 2004. Genomic organization, biology, and diagnosis of Taura syndrome virus and yellowhead virus of penaeid shrimp. *Adv. Virus Res.* **63**:353–421.