Tanay virus, a new species of virus isolated from mosquitoes in the Philippines

Takeshi Nabeshima,1 Shingo Inoue,1 Kenta Okamoto,1 Guillermo Posadas-Herrera,1 Fuxun Yu,1 Leo Uchida,1 Akitoyo Ichinose,2 Miako Sakaguchi,2 Toshikhiro Sunahara,3 Corazon C. Buerano,1,4 Florencio P. Tadena,4 Ildefonso B. Orbita,4 Filipinas F. Natividad4 and Kouichi Morita1

Correspondence
Kouichi Morita
moritak@nagasaki-u.ac.jp

1Department of Virology, Institute of Tropical Medicine, Nagasaki University, Sakamoto-machi 1-12-4, Nagasaki City 852-8523, Japan
2Central Laboratory, Institute of Tropical Medicine, Nagasaki University, Nagasaki University, Sakamoto-machi 1-12-4, Nagasaki City 852-8523, Japan
3Department of Vector Ecology and Environment, Institute of Tropical Medicine, Nagasaki University, Sakamoto-machi 1-12-4, Nagasaki City 852-8523, Japan
4Research and Biotechnology, St. Luke’s Medical Center, 279 E. Rodriguez Sr. Blvd, Quezon City 1102, Philippines

In 2005, we isolated a new species of virus from mosquitoes in the Philippines. The virion was elliptical in shape and had a short single projection. The virus was named Tanay virus (TANAV) after the locality in which it was found. TANAV genomic RNA was a 9562 nt + poly-A positive strand, and polycistronic. The longest ORF contained putative RNA-dependent RNA polymerase (RdRP); however, conserved short motifs in the RdRP were permuted. TANAV was phylogenetically close to Negevirus, a recently proposed taxon of viruses isolated from haemophagic insects, and to some plant viruses, such as citrus leprosis virus C, hibiscus green spot virus and blueberry necrotic ring blotch virus. In this paper, we describe TANAV and the permuted structure of its RdRP, and discuss its phylogeny together with those of plant viruses and negevirus.

In 2005, as part of our tropical arbovirus surveillance, we attempted to isolate mosquito-borne viruses in the Philippines. Mosquitoes were collected from a pig farm in Tanay, Rizal province, Luzon Island. Collected mosquitoes were Culex spp. and Armigeres spp. Mosquito pools were homogenized and centrifuged. The supernatants were passed through 0.22 μm filter and inoculated onto monolayer cultures of Aedes albopictus C6/36 cells, which were then incubated at 28 °C. Four samples (11-2 from Culex quinquefasciatus, 11-3 from Culex spp., 11-4 and 11-5 from Armigeres spp.) caused cytopathic effect to C6/36 cells. We attempted to amplify the viral genome by PCR using consensus primers of Flavivirus, Alphavirus and Bunyavirus, but were unsuccessful (data not shown).

To observe the virion, we conducted plaque purification three times, negatively stained the concentrated virus samples with 1.5 % uranyl acetate, and then examined them at 80 kV using a JEM-1230 transmission electron microscope (JEOL). We found elliptical particles (~50 nm) with a short single projection (Fig. 1a). Some particles seemed to be attached to each other through their projection.

To identify the unknown virus, RNA was extracted from the second passage infected culture fluids with TRIzol LS reagent (Invitrogen), and the RNA sequence was read comprehensively by a Roche 454 GS Junior pyrosequencer and a 3100-Avant Genetic Analyzer (Applied Biosystems). Reverse transcription-PCR, 3′-RACE and 5′-RACE for sequencing by the Sanger method were conducted as described previously (Nga et al., 2011). Data from pyrosequencing were assembled by GS De Novo Assembler (Roche).

From the four samples (11-2, 11-3, 11-4 and 11-5), almost identical sequences of RNA (Table S1 available in the online Supplementary Material) were obtained. The RNA had 9562 nt + poly-A and seemed to be polycistronic (Fig. 2a). We named the virus having this RNA as Tanay virus (TANAV). To confirm translation of the ORFs, we conducted an liquid chromatography/mass spectrometry
(LC/MS) analysis with liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) (Nano Frontier eLD; Hitachi) as described in a previous paper (Okamoto et al., 2010). LC/MS data were converted into the Mascot format by using data processing software (Hitachi High-Technologies), and analysed subsequently with MASCOT version 2.2 (Matrix Science KK; Perkins et al., 1999). LC/MS detected the peptides translated from the longest hypothetical ORF (ORF1) (RLLQLFSPPYRL, RITNTLSTLDYYQRN and KINADSAMKD) and ORF3 (KTAGRPVKVTEIRV, KFTTFKT), both shown in Fig. 2a.

To search for homologous sequences, BLAST+ version 2.2.28 (Camacho et al., 2009) was used. Dialign-tx v1.0.2 was used for multiple sequence alignment (Subramanian et al., 2008) with other viral genomes and proteins (Table S2). The aligned sequences were analysed by EMBOSs package 6.5.7 (Rice et al., 2000) and SEA VIEW 4.4.1 (Gouy et al., 2010). HMMER (http://hmmer.janelia.org) was used to detect specific protein motifs of the virus in the Pfam-A (Finn et al., 2014) database. The cut-off E value of HMMER was set to 1.0. ORF1 (Fig. 2a) was found to encode some RNA virus specific motifs: viral methyltransferase (Pfam: PF01660, 80–397 in ORF1), viral (Superfamily 1) RNA helicase1 (Pfam: PF01443, 1284–1554, E value 1.6 × 10⁻²²) and RNA-dependent RNA polymerase2 (RdRP_2, Pfam: PF00978, 1756–2196, E value 5.3 × 10⁻⁷⁰). However, in the RdRP of TANAV, a short motif, motif C [DX(4-5)D] and motif B [GX(2-3)TX(3)N; Fig. 2b], thus forming the permutation of the sequence motifs A, B and C from the canonical order A-B-C to C-A-B.

Similar permutation of motifs, giving the order C-A-B, was reported in some other RNA viruses. Fig. 2b shows the alignment of canonical (order A-B-C) and permuted RdRP_2 [TANAV, Dezidougou virus (DEZV), Santana virus (SANV) and grapevine virus Q (GVQ)] of RdRPs. The permuted RdRPs have a gap at the site corresponding to motif C in the canonical A-B-C type of RdRPs. DEZV and SANV, which were reported to be members of the recently proposed taxon Negevirus (Vasilakis et al., 2013), have a C-A-B type of RdRP_2. However, other members of the taxon Negevirus, namely Loreto virus, Negev virus (NEGV), Piura virus and Ngewotan virus (NWTV), have A-B-C type RdRP_2 (Figs 2b and 3b). In genus Marafivirus of the family Tymoviridae (ssRNA⁺ viruses), GVQ has a C-A-B type of RdRP_2 (Sabanadzovic et al., 2009), whereas the other members, grapevine rupestris vein-feathering virus and oat blue dwarf virus, have instead the A-B-C type of RdRP_2 (Figs 2b and 3b). RdRP of Birnaviridae, a family of dsRNA viruses, has the C-A-B type of RdRP (Birna_RdRp, Pfam: PF04197) (Duncan et al., 1991; Garriga et al., 2007; Gorbalenya et al., 2002; Pan et al., 2007) and the conserved sequence of motif C is ADN. Another dsRNA virus, the Drosophila A virus, also has a permuted C-A-B motif arrangement in the Birna_RdRp-like RdRp (Ambrose et al., 2009); however, the sequence of motif C is GDD. The family Permutotetraviridae is ssRNA⁺ and shares the Birna_RdRp-like C-A-B type of RdRP (Gorbalenya et al., 2002; Zeddam et al., 2010). This C-A-B type RdRp might have evolved independently in these various viruses, and these data provided an interesting example of convergent evolution of proteins.

Phylogenetic trees were reconstructed by the maximum likelihood (ML) method using PhyML 3.0.1 (Guindon et al., 2010) from a suitable area of alignment selected by trimAl 1.2rev59 (Capella-Gutiérrez et al., 2009). The substitution models were selected by Prottest 2.4 (Abascal et al., 2005).
Trees were drawn by NJPlot 2.4 (Perrière & Gouy, 1996). ML trees were reconstructed from protein sequence alignment of the ORF1 with viruses that shared homologous motifs (viral RNA helicase1 and RdRP_2; Figs 3a, b). The trees showed that TANAV formed a single clade with DEZV and SANV (Nege-CAB in Fig. 3), which have a permutation in RdRP homologous with that of TANAV. In the present phylogenetic trees of ORF1 (Fig. 3a, b), members of Negevirus formed a single clade with Citrus leprosis virus C (CiLV-C), Hibiscus green spot virus (HGSV) and Blueberry necrotic ring blotch virus (BNRBV) (CiLVp24 in Fig. 3); however, the negeviruses were separated into two subclades that could be discriminated by the translocation of motif C. From the order of the motifs, we named the clade with translocated motif C as Nege-CAB, and that with the non-translocated motif C as Nege-ABC.

The hypothetical ORF3 of negeviruses shared a conserved sequence similar to the p24 protein of CiLV-C (YP_654543),

![Fig. 2.](http://vir.sgmjournals.org) (a) ORF mapping in TANAV genomic RNA, strain 11-2. Black regions indicate hypothetical ORFs, and grey regions indicate conserved viral protein motifs and peptide sequences detected by LC-ESI/MS/MS. (b) Conserved motifs A, B and C in viral RdRP_2. This region corresponded to 1940–2106 aa in the TANAV ORF1. GenBank accession numbers are given in Table S2. (c) Predicted transmembrane region in ORF2 of Nege-CAB group of viruses. The transmembrane regions are shown in grey.

![Fig. 3.](http://vir.sgmjournals.org) ML phylogenetic trees. Bootstrap values greater than 600 are shown at the nodes. GenBank accession numbers are given in Table S2. (a) Tree built from ORF1, helicase region (amino acids 1288–1552 in the ORF1 of TANAV) by substitution model RtREV+I+G+F. (b) ORF1, RdRP region (1756–2215) by LG+I+G+F. (c) Complete ORF3 (1–215 in the ORF3 of TANAV) by LG+G+F.
p23 protein of HGSV (YP_004928121) and p24 protein of
BNRBV (YP_004901704); therefore, we named this group
CiLVp24 (Fig. S1). Negeviruses were also separated into two
clades, Nege-ABC and Nege-CAB, in ORF3, and TANAV
belonged to the latter clade (Fig. 3c).

Negevirus ORF2 is about 400 amino acids long and flanks
the 3'-end of ORF1; however, amino acid sequences were
not conserved between Nege-ABC and Nege-CAB, and no
other virus has a protein that is homologous with negevirus
ORF2 (Figs 2c and S2). TANAV ORF2 contained a DEZV-
or SANV-like conserved sequence and long insertion specific
for TANAV (amino acid positions 363–525 in ORF2). The
TANAV ORF2 contained 592 aa, and was distinctly longer
than ORF2 of other negeviruses (398–435 aa). The TANAV-
specific insert region contained the hypothetical transmem-
brane regions (Fig. 2c), which was predicted using SOSUI
(Hirokawa et al., 1998).

Clade CiLVp24 contained viruses with different genome
segment numbers. NEGV, a member of Nege-ABC, has a
non-segmented (single segment) RNA genome (Vasilakis
et al., 2013) and TANAV, belonging to Nege-CAB and phylogenetically distant from NEGV (Fig. 3), also has a
non-segmented RNA genome (Fig. 1b) determined by
electrophoresis of RNA from plaque purified virus sample
using a Formaldehyde-Free RNA Gel kit (Amresco). CiLV-
C and HGSV have, respectively, two (Locali-Fabris et al.,
2006) and three segments of RNA genome (Srinivasan
et al., 2005). BNRBV has a four-segmented RNA genome with
two helicase1 coding regions in segments 1 and 2; the
coding region on segment 2 is phylogenetically distant from
helicase1 of CiLV-C and HGSV. It is suspected that RNA
recombination between the ilarivirus (family Bromoviridae;
Fig. 3a)-like viral helicase1 domain and the CiLVp24 type
RdRP_2 domain occurred and the genome segment of
BNRBV is chimeric (Quito-Avila et al., 2013). This implies
that the segment number of clade CiLVp24 is not fixed, and
the number of genome segments can be a paraphyletic
character.

The newly proposed taxon Negevirus (Vasilakis et al., 2013)
corresponded to non-segmented genome viruses of the
clade CiLVp24 (Fig. 3). However, based on the phylogen-
etic trees of ORF1 and ORF3, groups Nege-ABC and Nege-
CAB are phylogenetically distant and appear as independ-
ent monophyletic clades (Fig. 3). In the RNA helicase1
region and ORF3, the Nege-ABC group is closer to CiLV-C
than to Nege-CAB. In addition, the virion of NWTv, a
member of the Nege-ABC group, is oval in shape (Vasilakis
et al., 2013); however, the virion of TANAV, a Nege-CAB
member, is elliptical with a projection (Fig. 1a). The two
species are morphologically distinct, and therefore the
phylogeny of this group should be reviewed.

Negevirus was isolated from haemophagic insects, mos-
quitos and phlebotomine sand flies, and it is suspected that
vertical infection, i.e. infection via egg production, is more
efficient than oral infection as the natural NEGV infection
route, because a high titre of NEGV (10^7–10^9 p.f.u. ml^{-1})
was required to infect Aedes mosquitoes orally from a blood
meal (Vasilakis et al., 2013). However, NEGV was isolated
from different genera of mosquitoes, Anopheles and Culicexp
(Vasilakis et al., 2013). TANAV was also isolated from two
different genera of mosquitoes, Culex and Armigeres, at the
same time and in the same location, and the RNA sequences
were almost identical between isolates (Table S1), suggest-
ing that TANAV is transmitted frequently between different
species of mosquito by some non-vertical route. CiLV-C,
HGSV and BNRBV were isolated from plants (Locali-Fabris
et al., 2006; Quito-Avila et al., 2013; Srinivasan et al., 2005)
and CiLV-C is transmitted by mites (Nunes et al., 2011).
Genomes of some insects, namely Drosophila rhopaloo, Bombus
terrestris and Aedes aegypti, contained endogenous
virus-like sequences similar to CiLV-C, HGSV and BNRBV
(Cui & Holmes, 2012). These data implied that some
CiLVp24 viruses are arthropod-borne plant viruses.

TANAV is the first isolate of the group Nege-CAB from
Asia. Another member of the Nege-CAB, DEZV, was
isolated from Ivory Coast in 1987, and SANV was isolated
from Brazil in 1992 (Vasilakis et al., 2013), indicating that
the group Nege-CAB is distributed worldwide. It
is suspected that many unknown members of the clade
CiLVp24 still exist. Although members of this clade were
found recently, only a few species have been identified
(Localii-Fabris et al., 2006; Quito-Avila et al., 2013;
Srinivasan et al., 2005; Vasilakis et al., 2013). Therefore,
it is necessary to find more species of this clade to further
understand this group.

Acknowledgements

We would like to thank the University of Rizal System in Tanay, Rizal,
Philippines, for allowing us to collect mosquitoes in their vicinity.
This work was supported by the Grant-in-Aid for Scientific Research
(KAKENHI) nos 18659131 and 25305015, the 21st century COE Program,
MEXT Japan, the Global COE program, MEXT, Japan, JSPS
Core University Program, the Health Labour Sciences Research Grant
and the Program of Founding Research Centers for Emerging and
Reemerging Infectious Diseases, IST, Japan. Also, it was partially
supported by a research grant from Research and Biotechnology, St.
Luke’s Medical Center, Quezon City, Philippines (01-017).

References


& Johnson, K. N. (2009). Drosophila A virus is an unusual RNA virus
with a T=3 icosahedral core and permuted RNA-dependent RNA

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J.,
Bealer, K. & Madden, T. L. (2009). BLAST+: architecture and appli-
cations. BMC Bioinformatics 10, 421.

trimAl: a tool for automated alignment trimming in large-scale


