Characterisation of the genomic sequence of an iflavirus, a protoambidensovirus, and of three microviruses detected in mosquitoes (*Aedes albopictus* and *Culex quinquefasciatus*)

Authors

Sarah François[a,b], Aymeric Antoine-Lorquin[a], Doriane Mutuel[a], Patrick Makoundou[c], Marco Perriat-Sanguinet[c], Sandra Unal[c], Hélène Sobry[a], Anne-Sophie Gosselin-Grenet[a], Mylène Ogliastro[a], Mathieu Sicard[c], Mylène Weill[c], Célestine Atyame[d] & Christophe Boëte[c#]

Affiliations

[a] DGIMI, Univ Montpellier, INRAE, Montpellier, France
[b] Department of Biology, University of Oxford, South Park Road, Oxford, OX1 3SY, UK
[c] ISEM, Univ Montpellier, CNRS, IRD, EPHE, Montpellier 34095, France
[d] Université de La Réunion, UMR PIMIT (Processus Infectieux en Milieu Insulaire Tropical) CNRS 9192, INSERM 1187, IRD 249, Saint-Denis, île de La Réunion, France.

Running title

insect-viruses and phages identified in mosquitoes

Corresponding author's email address

#Address correspondence to Christophe Boëte, christophe.boete@umontpellier.fr

Abstract

We report the complete CDS of five viruses: an iflavirus, a protoambidensovirus and three microviruses), which were detected by viromics surveillance of mosquitoes (*Aedes albopictus* and *Culex quinquefasciatus*) from the Réunion Island. We detected the protoambidensovirus, which belongs to a clade previously reported only in *C. pipiens*, in *A. albopictus*.

Announcement

Our knowledge of mosquito viruses remains limited, although they can modulate the transmission of arboviruses **(1)**. We processed 13 *Aedes albopictus* and 9 *Culex quinquefasciatus* larvae samples collected in La Réunion Island in 2019. Morphological characteristics were used for species identification, and larvae were pooled by location. Samples weighted between 0.5g and 1.49g, and stored at -80°C. Viromes were obtained as described in **(2)**. We followed manufacturers' instructions and default parameters except where otherwise noted. Samples were homogenised by a bead-beater, filtered through a 0.45 µm filter and centrifuged at 140,000 g for 2.5 hours. The pellets were resuspended, digested by DNaseI and RNaseA incubation at 37°C for 1.5 h. RNA and DNA were extracted using a NucleoSpin kit. Reverse transcription was performed by SuperScript III, cDNAs were purified by a QIAquick PCR Purification Kit and dsDNA synthetised by Klenow DNA polymerase I. DNA was amplified

by random PCR, and PCR products were purified using a NucleoSpin gel and PCR clean-up kit. Libraries were prepared using a NEBNext Ultra DNA PCR free with Ilumina adapter kit without fragmentation step, and sequenced on an Illumina HiSeq3000 in 2x150bp paired-end mode.

Adaptors were removed and reads were filtered for quality (q30 quality and read length >45 nt) using cutadapt 2.19 **(3)**. 10,150,317 reads were assembled into contigs by MEGAHIT 1.2.9 (4). Taxonomic assignment was achieved using DIAMOND 0.9.30 with against the NCBI nr protein database **(5)**. Genome coverage was assessed by mapping using Bowtie2 3.5.1 (end–to-end very-sensitive) **(6)**. Open Reading Frames (ORFs) were identified using ORF finder (length cutoff >400 nt) on Geneious prime 2024.0.5 **(7)**, and were annotated by blastn query-centered alignment against RefSeq viral database on October 2024. Genome circularisation was performed using Simple-Circularise (overlap >20nt).

We reconstructed the complete circular genomes of three bacteriophages species from the *Microviridae* family (*Gokushovirinae* subfamily) (PQ041302, PQ041303 and PQ041304). They share 54.5% to 75.0% of amino acid identity in major capsid and replication-associated proteins with their closest relatives, which were isolated from water environments (**Table 1**). There is currently no species demarcation criteria based on genome similarity defined by the ICTV for the *Microviridae* family **(8)**. Their high abundance in *Aedes* samples suggests than they may infect its microbiota, although their presence could also be due to environmental or diet contamination.

We report a complete CDS of a lineage of *XiangYun picorna-like virus 4* species (*Iflaviridae*, *Iflavirus*) (**Table 1**, PQ041301). It shares a whole polyprotein pairwise identity of 98.5% with a lineage discovered in *Culex theileri* from Yunnan, China (OL700176) **(9)**. This lineage clustered in a monophyletic clade containing only viral taxa isolated from dipteran species (mosquitoes and true flies), indicating that it is likely specific of dipterans (**10.6084/m9.figshare.28565888**). Further work is needed to determine the prevalence of XiangYun picorna-like virus 4 in natural populations of mosquitoes and its impact on mosquito fitness.

We report the complete CDS of a distant lineage of the mosquito infecting *Dipteran protoambidensovirus 1* species (*Parvoviridae*, *Densovirinae*, *Protoambidensovirus*, **Figure 1**, PQ041300). It shares 89.0% genome-wide nucleotide identity to its closest relative, Culex pipiens densovirus (FJ810126.1) (**Table 1**) **(10)**. This lineage belongs to the CpDV-3 clade which was previously only represented by sequences collected from Beijing (China) samples in 2003 **(11)**. Reported CpDV host range currently only includes *Culex pipiens* **(12,13)**, its broadening to *A. albopictus* warrants further confirmation.

Data availability

The genomic sequences of the five full-length viral genomes or CDS have been deposited at GenBank under the accession numbers PQ041300 to PQ041304. High-throughput sequencing reads were deposited in SRA under the accession no. SRR29133481 to SRR29133504 under PRJNA1114772 BioProject.

Acknowledgments

We thank Yann Gomard, Cyrille Lebon and Sarah Hafsah for their help in collecting samples in the field. This work was funded by the French ANSES PNR EST programme (project 2018/1/183, “DENSOTOOL", 2019–2023).

References

1. Olmo RP, Todjro YMH, Aguiar ERGR, de Almeida JPP, Ferreira FV, Armache JN, et al. Mosquito vector competence for dengue is modulated by insect-specific viruses. Nat Microbiol. 2023 Jan;8(1):135–49.

2. François S, Filloux D, Fernandez E, Ogliastro M, Roumagnac P. Viral Metagenomics Approaches for High-Resolution Screening of Multiplexed Arthropod and Plant Viral Communities. In: Pantaleo V, Chiumenti M, editors. Viral Metagenomics [Internet]. New York, NY: Springer New York; 2018 [cited 2024 Jan 29]. p. 77–95. (Methods in Molecular Biology; vol. 1746). Available from: A New Prevalent Densovirus Discovered in Acari

3. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011 May 2;17(1):10–2.

4. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015 May 15;31(10):1674–6.

5. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015 Jan;12(1):59–60.

6. Langmead B. Aligning Short Sequencing Reads with Bowtie. Current Protocols in Bioinformatics. 2010;32(1):11.7.1-11.7.14.

7. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012 Jun 15;28(12):1647–9.

8. Kirchberger PC, Ochman H. Microviruses: A World Beyond phiX174. Annual Review of Virology. 2023 Sep 29;10(Volume 10, 2023):99–118.

9. Feng Y, Gou QY, Yang WH, Wu WC, Wang J, Holmes EC, et al. A time-series meta-transcriptomic analysis reveals the seasonal, host, and gender structure of mosquito viromes. Virus Evol. 2022;8(1):veac006.

10. Cotmore SF, Agbandje-McKenna M, Chiorini JA, Mukha DV, Pintel DJ, Qiu J, et al. The family Parvoviridae. Arch Virol. 2014 May 1;159(5):1239–47.

11. Altinli M, Lequime S, Courcelle M, François S, Justy F, Gosselin-Grenet AS, et al. Evolution and phylogeography of *Culex pipiens* densovirus. Virus Evolution. 2019 Jul 1;5(2):vez053.

12. Altinli M, Lequime S, Courcelle M, François S, Justy F, Gosselin-Grenet AS, et al. Evolution and phylogeography of *Culex pipiens* densovirus. Virus Evolution. 2019 Jul 1;5(2):vez053.

13. Altinli M, Lequime S, Atyame C, Justy F, Weill M, Sicard M. Wolbachia modulates prevalence and viral load of *Culex pipiens* densoviruses in natural populations. Molecular Ecology. 2020;29(20):4000–13.

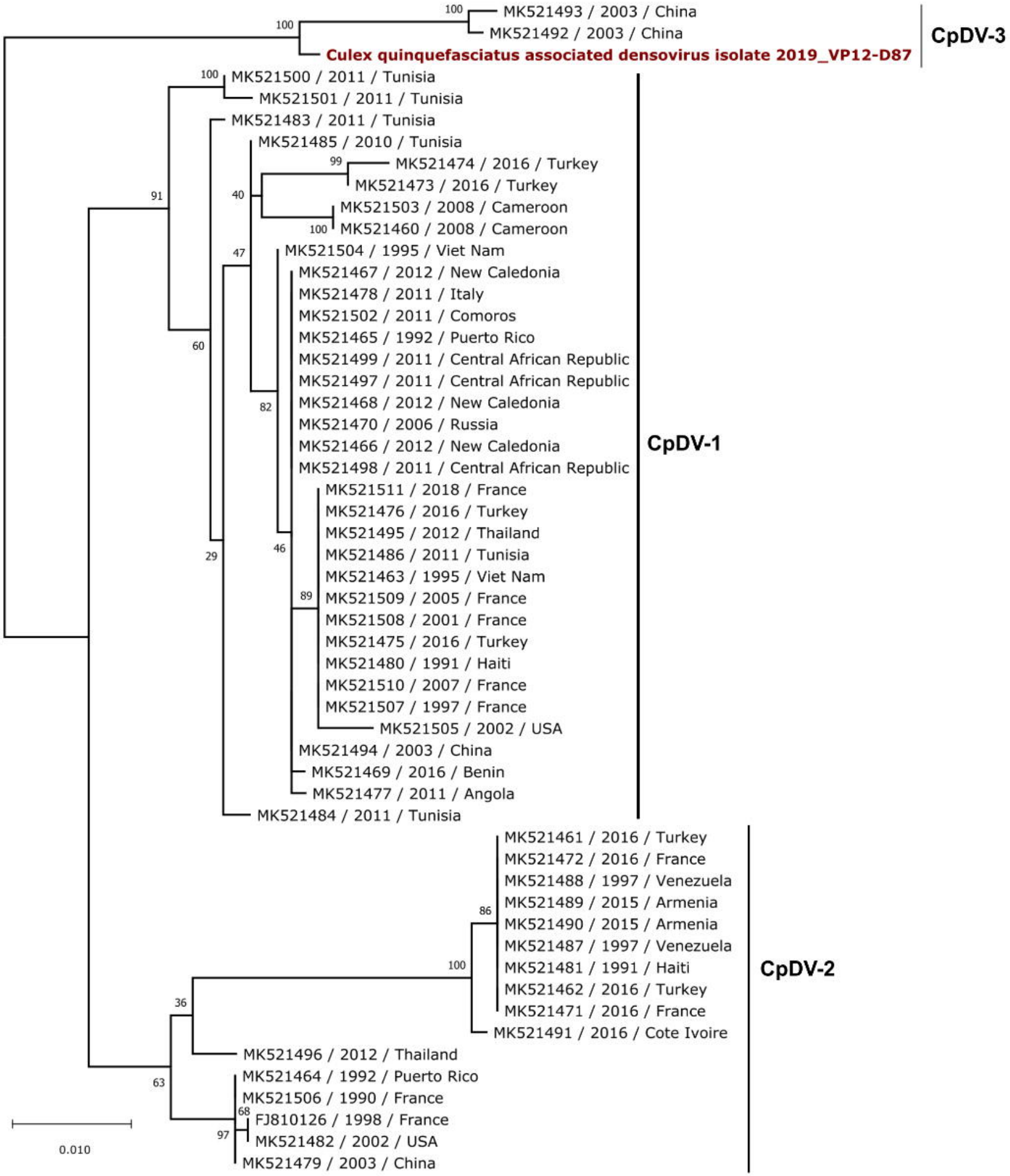
Figure legend

Table 1 Information on the five viruses reconstructed from mosquito viromic data

Figure 1 Maximum likelihood phylogenetic tree based on the NS1 nucleotide sequences of 52 *Dipteran protoambidensovirus 1* (common virus name: Culex pipiens densovirus (CpDV)) sequences. The alignment of 892 nucleotides in length was produced by MAFFT v7.450 using the G-INS-i algorithm. The tree is mid-point rooted. Bootstrap values (100 replicates) are indicated at each node. Scale bar corresponds to nucleotide substitutions per site. The CpDV sequence obtained from our sample is coloured in red.

Table 1 Information on the five viruses reconstructed from mosquito viromic data

Virus	GenBank accession number	Genome		Coverage			Putative proteins			Closest identified relatives			
		size (nt)	%GC	average	number of reads	sample	name	size (nt)	size (aa)	virus name	accession number	pairwise identity (aa)	source
						Cxqfts virome sample 3 (D87)				Culex pipiens densovirus	FJ810126		Culex pipiens
Culex quinquefasciatus associated densovirus, isolate 2019_VP12-D87	PQ041300	5 459	36,7	12 521	501 573	SRR29133503	NS1	687	229			94,7	
							NS2	372	214			92,2	
							NS3	642	124			42,4	
							VP	2 208	736			90,5	
						Cxqfts virome sample 1 (D85)							
Culex quinquefasciatus associated iflavirus, isolate 2019_VP12-D85	PQ041301	10 074	36,9	2 777	194 020	SRR29133486	polyprotein	9 231	3 077	XiangYun picorna-like virus 4	OL700176	98,5	Culex theileri
						Aedes virome sample 11 (D83)							
Aedes albopictus associated microvirus 1, isolate 2019_VP12-D83	PQ041302	4 365	43,9	243	8 198	SRR29133492	major capsid protein	1 650	550	Microvirus sp. isolate BS1_474	MT309944	75,0	wastewater
							DNA pilot protein	678	226	Microviridae sp. isolate CD-SXD-WGA-1-k141_518849	OR998997	43,3	bacteria
							internal scaffolding protein	429	143	Microvirus sp. isolate BS1_462	MT309947	54,2	wastewater
							replication initiator protein	1 023	341	Arizlama microvirus	MW697615	54,9	fresh water
						Aedes virome sample 3 (D75)							
Aedes albopictus associated microvirus 2, isolate 2019_VP12-D75	PQ041303	4 387	46,3	62	2 200	SRR29133501	major capsid protein	1 635	545	Microviridae sp. isolate SD_MC_72	MH572427	67,3	Ciona robusta
							DNA pilot protein	801	267	Microvirus sp. isolate LSGA_102419	OR349779	23,4	mollusc
							internal scaffolding protein	474	158	Microviridae sp. isolate ctda862	MH616902	65,8	fish
							replication initiator protein	747	249	Microvirus sp. isolate BS1_438	MT309954	74,7	wastewater
						Aedes virome sample 10 (D82)							
Aedes albopictus associated microvirus 3, isolate 2019_VP12-D82	PQ041304	4 813	50,5	197	6 618	SRR29133485	major capsid protein	1 653	551	Microviridae sp. isolate ctcc708	MH617066	73,4	fish
							DNA pilot protein	795	265	Microviridae sp. isolate ctji855	MH617641	37,6	fish
							internal scaffolding protein	444	148	Microviridae sp. isolate ctcc708	MH617066	61,9	fish
							replication initiator protein	783	261	Microvirus sp. isolate 1712115_622	MT310309	68,6	wastewater


MK521493 / 2003 / China
MK521492 / 2003 / China
Culex quinquefasciatus associated densovirus isolate 2019_VP12-D87
CpDV-3
MK521500 / 2011 / Tunisia
MK521501 / 2011 / Tunisia
MK521483 / 2011 / Tunisia
MK521485 / 2010 / Tunisia
MK521474 / 2016 / Turkey
MK521473 / 2016 / Turkey
MK521503 / 2008 / Cameroon
MK521460 / 2008 / Cameroon
MK521504 / 1995 / Viet Nam
MK521467 / 2012 / New Caledonia
MK521478 / 2011 / Italy
MK521502 / 2011 / Comoros
MK521465 / 1992 / Puerto Rico
MK521499 / 2011 / Central African Republic
MK521497 / 2011 / Central African Republic
MK521468 / 2012 / New Caledonia
MK521470 / 2006 / Russia
MK521466 / 2012 / New Caledonia
MK521498 / 2011 / Central African Republic
MK521511 / 2018 / France
MK521476 / 2016 / Turkey
MK521495 / 2012 / Thailand
MK521486 / 2011 / Tunisia
MK521463 / 1995 / Viet Nam
MK521509 / 2005 / France
MK521508 / 2001 / France
MK521475 / 2016 / Turkey
MK521480 / 1991 / Haiti
MK521510 / 2007 / France
MK521507 / 1997 / France
MK521505 / 2002 / USA
MK521494 / 2003 / China
MK521469 / 2016 / Benin
MK521477 / 2011 / Angola
MK521484 / 2011 / Tunisia
CpDV-1
MK521461 / 2016 / Turkey
MK521472 / 2016 / France
MK521488 / 1997 / Venezuela
MK521489 / 2015 / Armenia
MK521490 / 2015 / Armenia
MK521487 / 1997 / Venezuela
MK521481 / 1991 / Haiti
MK521462 / 2016 / Turkey
MK521471 / 2016 / France
MK521491 / 2016 / Cote Ivoire
MK521496 / 2012 / Thailand
MK521464 / 1992 / Puerto Rico
MK521506 / 1990 / France
FJ810126 / 1998 / France
MK521482 / 2002 / USA
MK521479 / 2003 / China
CpDV-2
0.010