

RESEARCH ARTICLE

A Metagenomic Survey of Viral Abundance and Diversity in Mosquitoes from Hubei Province

Chenyan Shi^{1,2}, Yi Liu¹, Xiaomin Hu¹, Jinfeng Xiong³, Bo Zhang¹, Zhiming Yuan^{1*}

1 Key Laboratory of Agricultural and Environmental Microbiology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China, **2** University of Chinese Academy of Sciences, Beijing, China, **3** Hubei Disease Control and Prevention Center, Wuhan, China

* yzm@wh.iov.cn



OPEN ACCESS

Citation: Shi C, Liu Y, Hu X, Xiong J, Zhang B, Yuan Z (2015) A Metagenomic Survey of Viral Abundance and Diversity in Mosquitoes from Hubei Province. PLoS ONE 10(6): e0129845. doi:10.1371/journal.pone.0129845

Academic Editor: Houssam Attoui, The Pirbright Institute, UNITED KINGDOM

Received: December 9, 2014

Accepted: May 13, 2015

Published: June 1, 2015

Copyright: © 2015 Shi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: This work was supported by Major State Basic Research Development Program (2013FY113500), Important National Science & Technology Specific Projects (2012ZX10004219) and Wuhan Key Laboratory on Emerging Infectious Diseases and Biosafety. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Mosquitoes as one of the most common but important vectors have the potential to transmit or acquire a lot of viruses through biting, however viral flora in mosquitoes and its impact on mosquito-borne disease transmission has not been well investigated and evaluated. In this study, the metagenomic technique has been successfully employed in analyzing the abundance and diversity of viral community in three mosquito samples from Hubei, China. Among 92,304 reads produced through a run with 454 GS FLX system, 39% have high similarities with viral sequences belonging to identified bacterial, fungal, animal, plant and insect viruses, and 0.02% were classed into unidentified viral sequences, demonstrating high abundance and diversity of viruses in mosquitoes. Furthermore, two novel viruses in subfamily Densovirinae and family Dicistroviridae were identified, and six *torque tenosus virus 1* in family Anelloviridae, three porcine parvoviruses in subfamily Parvovirinae and a *Culex tritaeniorhynchus* rhabdovirus in Family Rhabdoviridae were preliminarily characterized. The viral metagenomic analysis offered us a deep insight into the viral population of mosquito which played an important role in viral initiative or passive transmission and evolution during the process.

Introduction

Mosquitoes are important insect vectors to transmit amount of known pathogenic viruses in Family Flavivirus and Alphavirus, such as Dengue virus (DENV), West Nile virus (WNV) and Chikungunya virus (CHIKV), which threaten the public health and gain considerable attention worldwide [1, 2]. During the extrinsic incubation period of transmission cycle, mosquitoes are infected through ingestion of viremic blood from vertebrate host, then the viruses replicate and deposit into apical cavities of acinar cells, inoculating viruses into a new host upon refeeding [3]. In the infection process, viruses are subject to many bottlenecks such as evasion of the host and mosquito immune system which produce a certain evolutionary pressure on viruses leading to their mutation [3]. Since the hosts that mosquitoes feed on are highly diverse including

human, livestock, mammals, birds and plants [4], mosquitoes have access to amount of viruses or even novel viruses and are likely to be a viral reservoir or intermediate. In addition, some viruses can also be mechanically transmitted from one host to another by contamination mouth-parts or body surface of mosquitoes without infection or replication [3]. So surveillance for viral population in mosquitoes is critical for effective prediction of future vector-borne virus epidemic [5].

Hitherto, viral molecular detection of mosquitoes mainly focuses on greatly harmful and well-studied viruses by using degenerate PCR and real-time PCR [6, 7] which do not require *in vitro* replication and scarce serological or antigenic reagents [8]. These methods are efficient and specific, and can be widely used for identified mosquito-borne viral surveillance. However, the abundance and diversity of viruses as well as their effects on virus evolution and recombination have not been well investigated, and the risks of the identified and/or unidentified viruses on vector-born epidemic have not been evaluated. In consideration of the complex living environment of mosquitoes and their great virus carrying potential, it is significant to explore how abundant and diverse of viruses in mosquitoes.

With the development of second generation DNA sequencing technology, viral metagenomic analysis overcoming the limitation of traditional methods in viruses identification is considered to be an effective method not only for studying viral composition in various samples, such as respiratory tract [9, 10], feces [11, 12], blood [13], animal tissue [14], plants [15], *etc* or a specific biotope [16] but also for discovering novel viruses [17]. Previous researches demonstrated the biologically present of numerous viruses in a specific sample, and the cross transmission of viruses from one species to another which may initiate viral infectious diseases [18]. However, application of this technology in the characterization of viral flora in mosquito is rare yet. In previous study, viral metagenomic analysis was performed on mosquito samples collected in San Diego, USA to explore DNA virus population in them, reflecting a broad range of animal, plant, insect and bacterial viruses [5]. Besides, a novel densovirus infecting *Culex erythrothorax* was identified and animal viruses including anelloviruses, circoviruses, herpesviruses, poxviruses and papillomaviruses were detected using this technique.

In this study, the abundance and diversity of viruses in three mosquito samples from Hubei, China, were analyzed by metagenomic sequencing. The novel mosquito densovirus and Big Sioux River virus were identified, and six *torque teno sus virus 1*, three porcine parvoviruses and a *Culex tritaeniorhynchus* rhabdovirus were characterized. The data provides a preliminary baseline of DNA and RNA virus community in mosquitoes, as well as the animal and plant host they feed on. Highly diversity of viral flora in mosquitoes, as a widespread vector, underlines its function in virus transmission and evolution.

Results and Discussion

Overview

1,202 mosquitoes were collected at four locations in Hubei province during July 2013. Sample I was a collection of *Culex tritaeniorhynchus*. Sample II was *Anopheles sinensis*. Sample III was a mixture pool of *Armigeres subalbatus* and *Culex fatigans* (Table 1). The four species are the principal vectors of Japanese encephalitis virus (JEV) and *Plasmodium vivax* in Hubei.

The reads generated from the 454 High Throughput FLX short-gun sequencing were segregated into bins based upon barcode sequences and then the unique tags were trimmed from the ends. A total of 92,304 trimmed valid reads were generated after removing sequences shorter than 50 nt and the reads average length was 247 nt (Table 2). 39.2% viral sequences, identified by querying the viral database from NCBI running BLASTn and BLASTx at the E-value cutoff of 0.01, had high similarities to a variety of different viruses (Table 2), with 3.6% of

Table 1. Mosquito Samples used for metagenomic analysis.

Samples	Species	Numbers	Locations
Sample I	<i>Culex tritaeniorhynchus</i>	438	Hubei province ^a
sample II	<i>Anopheles sinensis</i>	424	Hubei province ^a
Sample III	<i>Armigeressubalbatus</i>	259	Hubei province ^a
	<i>Culex fatigans</i>	92	Hubei province ^a

^aCollection sites include Zhouji (N 30° 24', E 112° 48') and Yangshi (N 30° 21', E 112° 55') in Qianjiang city, Sancha town (N 30° 56', E 114° 03') and Dawu (N 31° 33', E 114° 08') in Xiaogan city.

doi:10.1371/journal.pone.0129845.t001

vertebrate virus, 88% of insect virus, 0.8% of plant virus, 1.87% of bacteriophage and 0.03% of mycovirus (Fig 1A) among them. However, the abundance and diversity of viral sequences were different in three samples, with the largest viral reads number of Sample III and highest viral diversity in Sample I. Insect virus occupied the most in Samples III, but viral reads of Sample I scatteredly distributed in each virus type, which may attribute to the wide distribution of *C. tritaeniorhynchus*, the main vector of JEV in China. Though the gross reads of Sample II was less comparatively, the viral reads accounted for 21% and the majority were vertebrate viruses including Anelloviridae and Parvovirinae which also found in other samples.

All viral reads were classified on family level (Fig 1B). Among these viruses, some were biologically or mechanically transmitted by mosquitoes or specifically infected mosquitoes, the other were probably sucked sporadically by mosquito from host reservoirs. Additionally, considerable unidentified sequences highly likely belonged to some unexplored novel virus that dissimilative with known virus within mosquitoes, needing further investigation.

Family Anelloviridae

Blast results showed that all three samples contained the viral sequences closely related to *torque tenos virus 1* (TTSuV1), Genus *Iotatorquevirus*, Family Anelloviridae. 8, 6 and 13 reads of each sample were assembled into one contigs with 745 nt, 339 nt, 484 nt separately. At nucleic acid level, contigs of Sample I and Sample III had high identity (nearly 99%) with isolated TTSuV1. In Sample II, contigs only showed 84% identity to known TTSuV1.

A pair of universal primers for TTSuV1 (Table 3) were used to amplify a 678 bp genomic fragment including the untranslated region (UTR), the complete non-structure protein 2 gene (ORF2) and 191 bp of the 5' end of capsid protein gene (ORF1) [19, 20] from the three samples. One sequence was identified in Sample I, three in Sample II and two in Sample III, which were named as TTSuV1 HB1, HB2-1/2/3 and HB3-1/2 respectively. These sequences have been deposited into GenBank (KR131713-KR131718). A neighbor-joining phylogenetic tree was constructed using the 6 amplified sequences and corresponding sequences of other 36 TTSuV1 stains in GenBank (Fig 2) and the identified TTSuV1 were closely related (> 90% identity) to known TTSuV1 isolated in China and subtyped into TTSuV1a, 1c and 1d [21]. Further analysis

Table 2. Basic Data of 454 FLX Sequencing.

Samples	Barcodes	No. of base	No. of reads	Ave. length (nt)	No. of viral reads	Percentage of viral reads
Sample I	A	12,416,751	51,123	242	7,298	14%
Sample II	B	190,921	823	231	170	21%
Sample III	C	10,786,900	40,359	267	28,717	71%
Total / Average		23,394,572	92,304	247	36,185	39%

doi:10.1371/journal.pone.0129845.t002

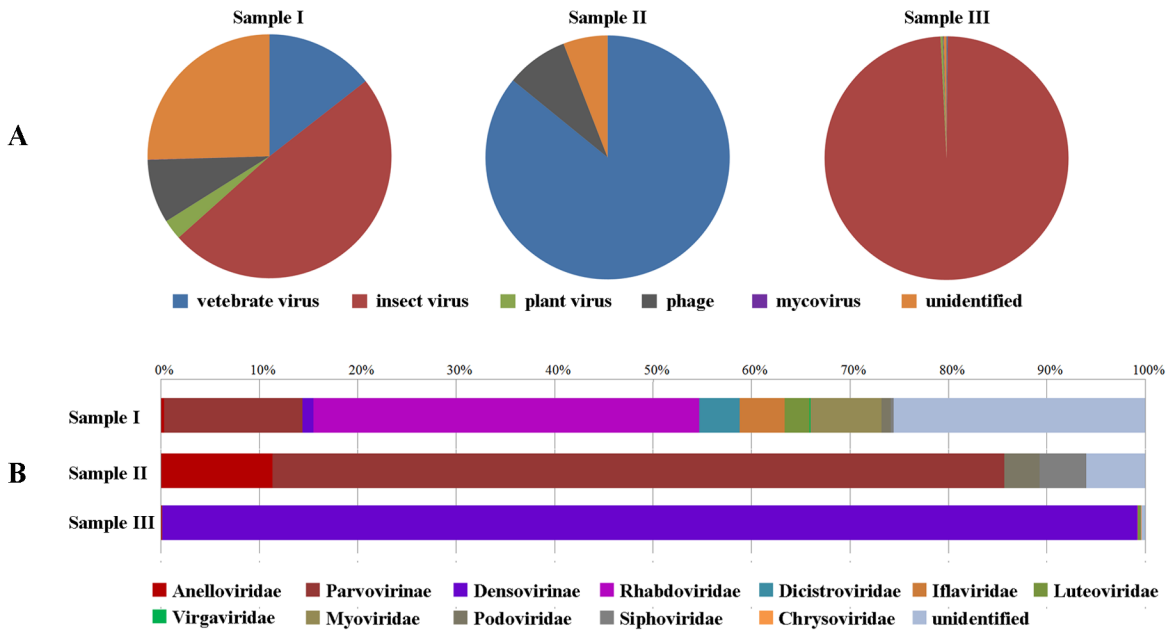


Fig 1. Taxonomic classification of viral sequences in three samples. (A) Viral sequences were classified by host type. Percentage of Mycovirus was too low which hardly can be seen from figure. (B) Viral sequences were classified on family level. The family with reads number less than 10 did not show in figure. Different host types and families were denoted with different colors.

doi:10.1371/journal.pone.0129845.g001

revealed that TTSuV1 HB1 and HB2-3 were in the same clade belonging to TTSuV1c with 96% identity to the strain TTV1Bj6-1. TTSuV1 HB2-1, HB2-2, and HB3-1 were classified into TTSuV1d with 93%, 90% and 97% similarity to strain TTV1cHK, Bj1-1 and Fj3 respectively. TTSuV1 HB3-2 had 97% identity to stain TTV1Bj3 in the subtype TTSuV1a.

Although the distribution of TTSuV was investigated among 14 provinces of China [21], TTSuV1a, 1c and 1d were firstly identified in mosquitoes from Hubei, suggesting the possible existence of the virus in pig population. The frequent variation of TTSuV had been noticed and it often co-infected with other viruses, especially the porcine circovirus, which posed potential threaten to pig herds [22].

Subfamily Parvovirinae, Family Parvoviridae

All three samples contained viral sequences belonging to Genus *Parvovirus*, subfamily Parvovirinae and the blast results suggested that these sequences matched most closely to porcine parvovirus (PPV). Sample I contained the majority of PPV sequences (991 reads), they assembled into 18 contigs (Table 4). Among them, seven contigs shared 99% nucleotide identity to PPV4, five were most closely related to PPV5 and six best matched with PPV6. 112 reads from sample II formed 6 contigs which were most closely to PPV4 (Table 4) and one contigs of Sample III comprised of 14 reads best matched with PPV5 (Table 4). PPV4 isolate WB-209CV (GenBank: JQ868714.1), PPV5 isolate IN273 clone 1 (GenBank: JX896319.1) and PPV6 isolate SC (GenBank: KF999684.1) were used as reference to map these contigs and the gaps were filled by PCR (Fig 3). The assembled contigs were aligned with other known PPVs and phylogenetic analysis results indicated all contigs have close relationship with the reference strains (Fig 4).

Porcine parvovirus (PPV) infection was always characterized by SMEDI (Stillbirth, Mummification, Embryonic Death and Infertility). The classical PPV (designated as PPV1) was first isolated in Germany and America in 1965 [23, 24] and has been proved wide distribution

Table 3. Primers used in this study.

Product names	Orientation	Sequence	Product size (bp)
I-PPV4 fragment1	F	ACAGTGTCTGAGAATCGCTCA	2022
	R	GTGACCGTGTTTCTGCATTGT	
I-PPV4 fragment2	F	GTGACCGTGTTTCTGCATTGT	1213
	R	GCTCATCACTTGATCCTCTG	
I-PPV4 fragment3	F	TGACAATAACTCTGCTACACTG	1060
	R	GTGTTCCGGTATCTGAGACCAT	
I-PPV5 fragment1	Ex-F ^a	TGCAATTGGAAAAGATTGACTC	2385
	Ex-R ^a	ACCTGTGTAGCGATGTCTCTG	
	In-F ^b	AGGTACAATAGGTGAAGCCTG	1955
	In-R ^b	CTTGATTGGTGATGTTCCCTCAT	
I-PPV5 fragment2	Ex-F	AGTACCTGTGAGACCAGACAA	2596
	Ex-R	GGTTCACCTCGGAGTATGT	
	In-F	ATCACTGCTGACTTCGATCAG	2386
	In-R	TTGCTTCCAGAACAGCTGGAT	
I-PPV6 fragment1	F	TATGTCTCAGACCTTCTGGAC	1667
	R	CAATCCTCTGGAGGAGTATCT	
I-PPV6 fragment2	Ex-F	TCACATCCAATACGGATATGTG	2266
	Ex-R	CAGCTCATAACACAGGTTCTGT	
	In-F	ATTTCTGAAGATGAAGTCGTGG	1942
	In-R	ATGTCTCGCAGAAGATTCATCA	
I-PPV6 fragment3	F	TCTTCTGCGAGACATGACATC	1708
	R	CCTATACATGGTATGTCTGACA	
II-PPV4 fragment1	F	CATGGCTAAACAGTGTCTGAG	637
	R	GGATCTCTCCAACAGGAACTA	
II-PPV4 fragment2	F	GCTGGGCAAGGAATGTCAATT	286
	R	TCTCAACCGGAAGTCTGTAA	
II-PPV4 fragment3	F	TGATCAACTTGACGAAGCAGC	427
	R	CGCTGAGAGAACACATCTTCT	
II-PPV4 fragment4	F	AGTGAACCCATACATTGTGCC	410
	R	TGACCTGGCATCGAATATCCT	
III-DNV fragment1	Ex-F	GGAGAGCGAGATCAGTAATGA	2145
	Ex-R	GTGGCCGACTGATGAGTTTG	
	In-F	TGCATTTATTGTTGGGAGCATG	1904
	In-R	TGATCCATGTTGGCGTTCTCT	
III-DNV fragment2	Ex-F	TGTGTAGAGTGGATCGAATAC	1955
	Ex-R	GCTATATGGATCATCGGAGGT	
	In-F	GGAAGGCATAACAATGCTGG	1797
	In-R	GGTGGACGTAGAAGGTGGAA	
I-BSRV fragment1	Ex-F	TCATTGTGTGTTCTCTGGTG	3025
	Ex-R	CAGGTGCTGTTGTTCTGTCAA	
	In-F	GTCATATGGCATCGATATTGGA	2597
	In-R	TCTTGCTCCAAGACTGGTCG	
TTSuV1	F	CGGGTTCAGGAGGCTCAAT	678
	R	TCTACGTCCCCTCTACGG	

^aPrimer pairs for the first round of nested-PCR

^bPrimer pairs for the second round of nested-PCR

doi:10.1371/journal.pone.0129845.t003

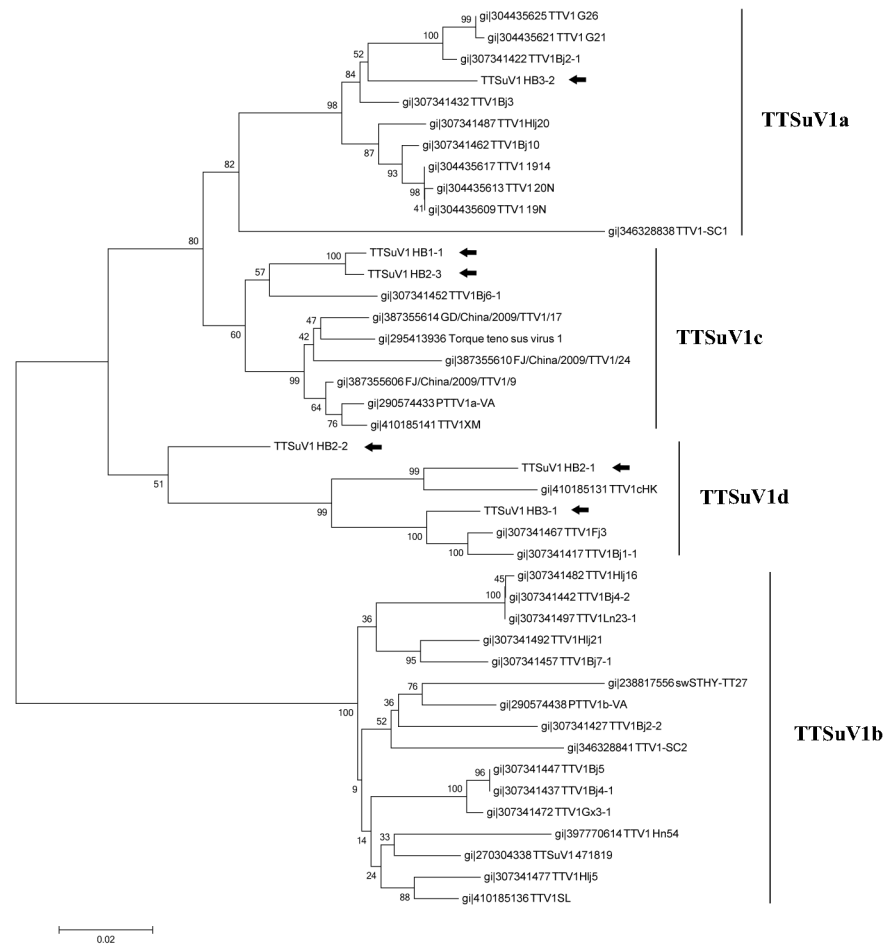


Fig 2. Neighbor joining phylogenetic tree of TTSuV1 based on 678 bp segment. 678 bp contained untranslated region (UTR), the complete NS2 gene (ORF2) and 191 bp of the 5' end of capsid protein gene (ORF1). Six TTSuV1 sequences characterized in this study were marked with black arrow.

doi:10.1371/journal.pone.0129845.g002

around the world especially in swine herds, causing economic loss for swine breeding industry. In recent years, a lot of novel PPVs had been identified using improving biology molecular technologies worldwide, but the classification and infection of these novel viruses were still unknown. The high mutation rate of capsid gene of PPV like RNA virus [25] and the evolution pressure of vaccine used currently might enhance the generation of novel PPVs. The increasing numbers of new PPVs discovered in different parts of the world indicated the endless evolution of PPV and suggested the vaccine produced from PPV strain decades years ago should be modified.

The presence of PPV and TTSuV1 in mosquitoes probably attributed to the contingent blood sucking of the mosquitoes from viremic pigs. Whether mosquito was an intermediate host to disseminate these porcine viruses or drew diseased blood coincidentally were needed to be verified in the subsequent study. Whatever, the mosquito control of piggens was a crucial and effective step to keep pigs health.

Table 4. Contigs Information of Porcine Parvovirus.

Sample	Contigs No.	No.of Assembled Reads	Length (nt)	Best match	Identity	E value
Sample I	1	298	2083	PPV4 WB-209CV	99%	0.0
Sample I	2	232	1323	PPV6 SC	99%	0.0
Sample I	3	54	1029	PPV4 WB-209CV	99%	0.0
Sample I	4	84	981	PPV5 IN273	99%	0.0
Sample I	5	14	972	PPV6 BJ	99%	0.0
Sample I	6	19	894	PPV6 BJ	100%	0.0
Sample I	7	16	854	PPV4 WB-209CV	99%	0.0
Sample I	8	165	806	PPV5 IN273	99%	0.0
Sample I	9	7	699	PPV4 WB-178SM	99%	0.0
Sample I	10	11	686	PPV5 IN273	99%	0.0
Sample I	11	8	655	PPV6 BJ	99%	0.0
Sample I	12	30	654	PPV6 BJ2	100%	0.0
Sample I	13	15	585	PPV5 IN273	99%	0.0
Sample I	14	14	540	PPV5 MI216	98%	0.0
Sample I	15	13	450	PPV6 JS	99%	0.0
Sample I	16	7	437	PPV4 WB-549HB	99%	0.0
Sample I	17	10	432	PPV6 SC	99%	0.0
Sample I	18	6	327	PPV4 WB-549HB	99%	2e-166
Sample II	1	32	1072	PPV4 WB-209CV	99%	0.0
Sample II	2	10	822	PPV6 SC	99%	0.0
Sample II	3	19	718	PPV4 WB-209CV	99%	0.0
Sample II	4	26	696	PPV4 F2-11SM	99%	0.0
Sample II	5	8	589	PPV4 WB-209CV	99%	0.0
Sample II	6	6	444	PPV4 WB-209CV	99%	0.0
Sample III	1	7	507	PPV5 IN273	99%	0.0

doi:10.1371/journal.pone.0129845.t004

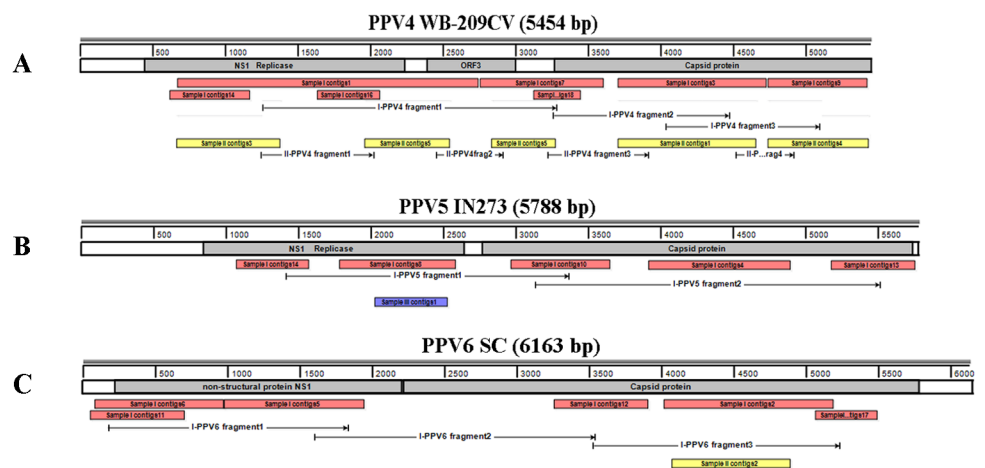


Fig 3. Alignment of PPV contigs in three samples. ORF of reference was labeled with grey strips. Contigs of three samples filled with red, yellow and purple respectively. Stroked arrow denoted the amplified fragments through PCR.

doi:10.1371/journal.pone.0129845.g003

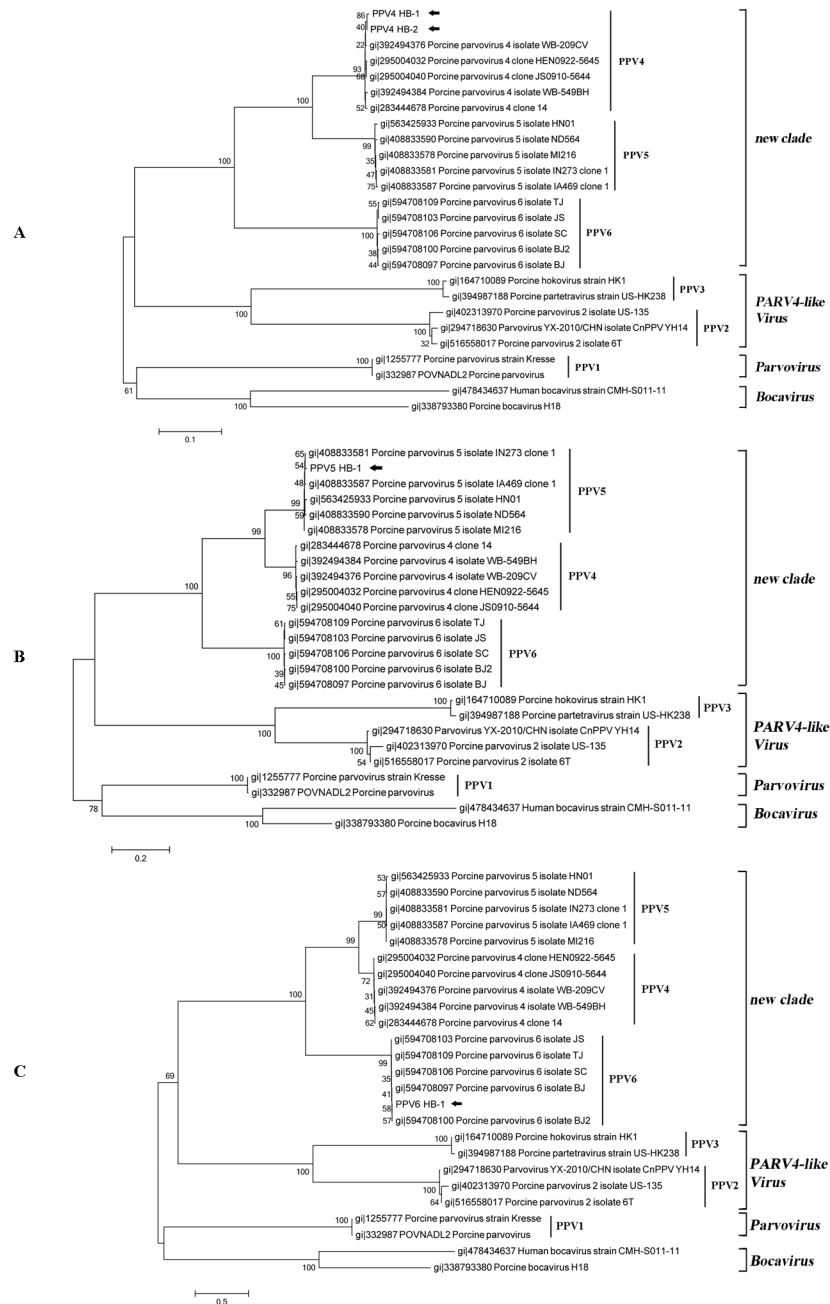


Fig 4. Neighbor joining phylogenetic tree of PPV based on nucleic acid of NS1. PPV sequences identified in this study were labeled with black arrows.

doi:10.1371/journal.pone.0129845.g004

Subfamily Densovirinae, Family Parvoviridae

Sample III contained the largest number of viral reads and majority of them had similarities with mosquito densovirus (DNV), genus *Brevidensovirus*, subfamily Densovirinae. More than 27,000 sequences were assembled into 58 contigs, of which the longest was 1084 nt. All contigs showed relative low identity (average identity was 85%) to known densovirus according to the BLASTn results. *Culex* densovirus 0507JS11 (GenBank: FJ805445.1) was used as a reference genome to align the 58 contigs which covered the whole reference genome (Fig 5). Four pairs of

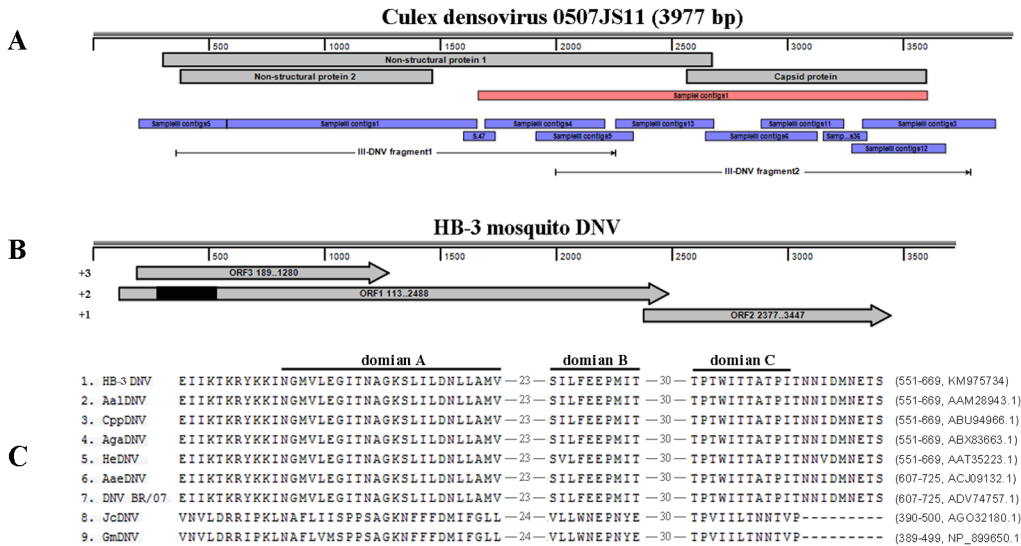


Fig 5. Amplification and sequence analysis of DNV in Sample III. (A) The mapping of DNV contigs in this study. It only presented the contigs used for assembly. ORF of reference was labeled with grey strips. Stroked arrow denoted the amplified fragments through PCR. (B) The ORF prediction of the sequenced novel DNV in Sample III. Black part (279 bp to 545 bp) of ORF1 was identified as Parvo_NS1 Superfamily. (C) 119 aa highly conserved region within NS1 of novel DNV like other DNVs. Domain A was the NTP-binding part. B and C were helicase domain.

doi:10.1371/journal.pone.0129845.g005

primers (Table 2) were designed referring to contigs nucleotide sequences and two nested PCRs were performed to amplify main genomic sequence. Two products, about 1791 bp and 1623 bp in length, were then cloned to vectors for sequencing, and a whole viral genome was assembled (KM975734), named as HB-3 mosquito DNV.

HB-3 mosquito DNV genome contained three ORFs through three different reading orders (Fig 5B). ORF3 encoded a protein of 363 amino acid (aa), having 73% similarity with non-structure protein 2 (NS2) of known DNVs. Translation product of the largest ORF1 with 791 aa showed 83% identity to NS1 of known DNVs. Besides, a highly conserved domain (279 bp to 545 bp) in ORF1, with the DNA helicase and ATPase activity, was identified as Parvo_NS1 Superfamily (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?RID=Z5BKE39M013&mode=all>). The right ORF2 encoded a product of 356 aa, sharing 82% identity to capsid protein of known DNVs. Multiple alignment of amino acid sequence revealed that NS1 of HB-3 mosquito DNV had two highly conserved regions like other members of family Parvoviridae [26]. One from aa 316 to 375 (GNHLHLIF-20-FILYCIYGI) was the highly conserved replication initiator motif [27], another between aa 551 and 699 contained the NTP-binding sequence and helicase domain [27–29] (Fig 5C). Function of the most highly conserved protein NS2 whose aa sequence was distinctly different from others was still unknown. Capsid protein of HB-3 mosquito DNV involved a characterized highly conserved sequence (RGTKRKR, aa 14 to 20) in the glycine-rich region which also has been found in AaeDNV, AalDNV, CppDNV and HeDNV [30]. Amino acid sequence of NS1 was used for exploring the relationship between HB-3 mosquito DNV and other DNVs (Fig 6A), indicating that the novel DNV was apparent different from other mosquito DNVs. However, the characterization of HB-3 DNV should be performed for well understanding its biological property and its potential application.

Besides, 77 DNV reads in Sample I was assembled into one contigs with 1949 nt (Fig 5A). It showed <84% similarity at nucleotide level to known densoviruses and located in the latter

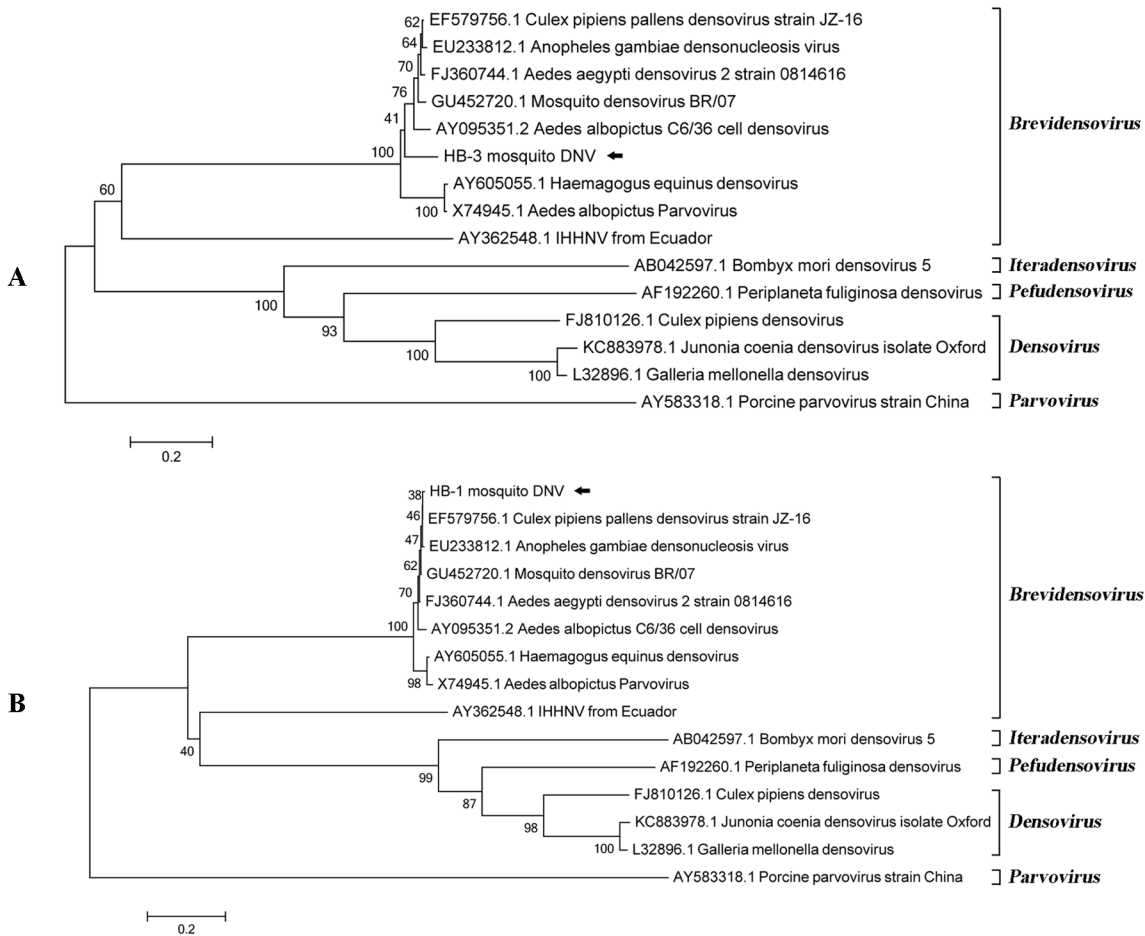


Fig 6. Phylogenetic analysis of DNV using neighbor joining method. (A) phylogenetic trees constructed based on aa sequence of putative ORF1 of HB-3 DNV and NS1 of other DNVs. (B) Partial nucleic acid sequence of Sample I contigs located in NS1 region was used for Phylogenetic analysis. DNVs identified in this study were marked with black arrows.

doi:10.1371/journal.pone.0129845.g006

half of DNV genome. The partial genesequences located in NS1 was used for phylogenetic analysis (Fig 6B), suggesting its close relationship with *Culex pipiens pallens* densovirus.

Densovirinae includes three genus *Densovirus*, *Iteravirus* and *Brevidensovirus*. Among them, *Brevidensovirus* have been identified in or isolated from various mosquito cell line and wild caught mosquitoes, demonstrating the extensive distribution of mosquito densovirus [30, 31]. The viral type species of the genus was *Aedes aegypti* DNV (AaeDNV) isolated from mosquito larvae [28], and some other DNVs were found in a chronically infected *A. albopictus* C6/36 cell line (AalDNV) [32], *C. pipiens* larvae (CpDNV) [33], and wild caught adult *C. pipiens pallens* (CppDNV) in China [30]. Besides one DNV isolated from shrimp (infectious hypodermal and hematopoietic necrosis virus, IHHNV) also showed high similarity to *Brevidensovirus* on genomic level [27].

The host ranges of mosquito DNV were narrow, but they can perpetuate through horizontal and vertical transmission in mosquitoes life cycle from infected adult to larvae [34]. Due to its high lethality to cell line and early instar larvae, as well as its persistent infection and transmissible viremia in late instar larvae and adult [31, 34] mosquito DNV could be used as viral agent for mosquito control.

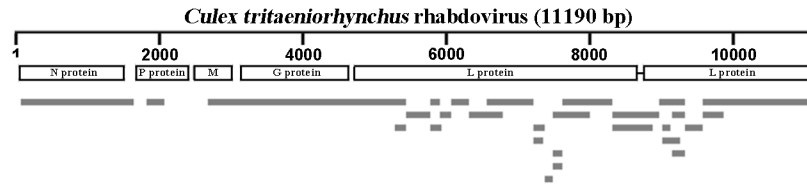


Fig 7. Alignment of CTRV contigs in Sample I. Grey strips were the contigs of CTRV in Sample I. Short line between two parts of L protein was the intron sequence.

doi:10.1371/journal.pone.0129845.g007

Family Rhabdoviridae and Dicistroviridae

The majority sequences in Sample I and several in Sample III matched to family Rhabdoviridae and all of them were most similar to *C. tritaeniorhynchus* rhabdovirus (CTRV), which has only been identified in Japan in 2011. Reads of Sample I formed 30 contigs with 98% identity to CTRV and covered the whole genome of CTRV (GenBank: AB604791.1) (Fig 7).

Six contigs, assembled from the reads of Sample I, were most closely related to two species: *Rhopalosiphum padi* virus (RhPV), Genus *Cripavirus* and Big Sioux River virus (BSRV), unclassified. On account that the BSRV had close relationship with RhPV [35] and only partial genomic sequences of BSRV could be found in GenBank, RhPV6 (GenBank: EU282007.1) was chosen to map the contigs (Fig 8A). Nested PCR was designed to amplify the gaps between the 4 contigs located in the gene of non-structure protein, denoted in Fig 8A. Furthermore, the 1473 bp sequence in nonstructural polyprotein gene of BSRV in GenBank was used for phylogenetic analysis with viruses identified in this study and other viruses in genus *Cripavirus* (Fig 8B). The result showed that newly identified sequences had closer relationship with BSRV in Family Dicistroviridae. HB-1 BSRV share 87% identity with BSRV1 and the E value was 0, which indicated likely to be a novel BSRV.

Rhabdoviruses had a wide range of hosts and had been identified in plants, vertebrate animals and insects [36]. However, no evidences showed their vertical transmission in plants and vertebrate animals, so insects as mechanical or biological vector played a part in their horizontal transmission [37, 38]. CTRV genome encoded 5 proteins like other rhabdoviruses, but its gene sequence of L protein, a RNA-dependent RNA polymerase, embraced an intron section which needed RNA splicing during transcription. This process was rare in negative single-

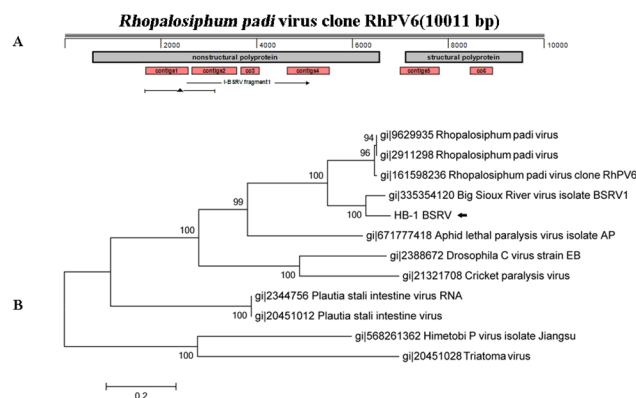


Fig 8. Alignment and phylogenetic analysis of contigs in Family Dicistroviridae of Sample I. (A) The mapping of contigs in Family Dicistroviridae of Sample I. ORF of reference was labeled with grey strips. Stroked arrow denoted the amplified fragments through PCR. Triangle glyph with a line denoted the region used for phylogenetic analysis. (B) 1473 bp of the non-structure protein gene was used for phylogenetic analysis. BSRV detected in Sample I was labeled with black arrows.

doi:10.1371/journal.pone.0129845.g008

strand RNA viruses and had never been found in other members of Family Rhabdoviridae [39]. In phylogenetic analysis, CTRV was closely related to genera *Vesiculovirus* and *Ephemerovirus* which mainly infected vertebrate animals [39].

Partial genomic sequence of BSRV was only identified in western honey bees of USA in 2011. It most closely related to RhPV on the nucleic acid level, but shared only 78% and 69% aa identity with NS polyprotein and capsid protein of RhPV respectively, which met the taxonomic rank of a new species on the level of aa divergence [35]. The presence of BSRV in mosquitoes from China may indicate the wide existence of the virus, and its genome sequence, infectivity and transmission pathway need to be characterized in the following study.

Materials and Methods

Mosquito collection

Our mosquitoes were collected on private land and the owner of the land gave permission to conduct the study on this site. Sample collection was conducted with the help of Hubei CDC and no protected species were sampled, so no specific permissions were required for these locations/activities.

The adult mosquitoes were collected from the cowsheds and pigpens located in different residential areas in Hubei province by using an electrical mosquito aspirator (Yalin, Zhejiang, China). The collected mosquitoes were transported to laboratory alive in cages and killed by freezing at -20°C for 30 min before classified into species with approximately 50 adult per pool. All mosquitoes were classed into three samples (I, II and III) for metagenomic analysis (Table 1).

Nucleic acid extraction and library construction

Pools of mosquitoes were homogenized using TissueLyser II (Qiagen) and 3 mm tungsten carbide beads (Qiagen) in 1.5 ml sterile plastic tubes and 1 ml viral transport medium (VTM) containing Hank's solution, 1% BSA, 15 $\mu\text{g}/\text{ml}$ Amphotericin, 100 unit/ml Penicillin and 50 unit/ml Streptomycin. Samples were centrifuged at $13,000 \times g$ for 30 min at 4°C , taking the supernatant for second centrifugation at the same conditions. Store the supernatant at -80°C until use.

120 μl supernatant was treated with Turbo DNA-free Kit (Ambion), Benzonase (Novagen) and RNase I (Fermentas) to digest unprotected nucleic acid. Nucleic acid packaged in viral capsids were extracted using QIAamp Viral RNA Mini Kit (Qiagen) without DNase digestion [11]. The extracts including both RNA and DNA viral sequences were used for reverse transcription with random primers containing short specific barcode sequences at 5' end by M-MLV Reverse Transcriptase (Invitrogen), following by DNA extension reaction with Klenow fragment polymerase (Fermentas) and the same primers. Subsequently, PCR amplification was performed with Primer STAR HS (TaKaRa) and reverse random primers. PCR products were analyzed by 1% agarose gel electrophoresis. Then DNA between 500 to 1000 bp was cut and purified with gel extraction kit (Omega) and sequenced using Roche 454 GS FLX system. Three samples used different primers with specific barcode sequences (barcode A, B, C respectively) which could be recognized in the following analysis.

Bioinformatics analysis and phylogenetic analysis

The reads of three samples were separated by barcode sequences and barcode sequences were trimmed. Trimmed reads that shorter than 50 nt were discarded and then the remainder were compared against GenBank database using both BLASTn and BLASTx with E value set as 0.01. Sequences were divided according to the best match. Viral de novo DNA sequences within one

species or genus were assembled using Newbler (Roche) [40] with the default parameters. Mapping of contigs used BLASTn. ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>) of NCBI was used for ORF prediction.

Multiple alignment of sequence was performed with ClustalW. All phylogenetic trees were performed with MEGA6 [41] using neighbor joining method and bootstrap set as 1000 replications.

PCR amplification and Sequencing

Primers used in the study were in Table 4. Amplification of the gaps between contigs were conducted using Primer STAR HS (TaKaRa) and reaction condition was 98°C for 5 min, followed by 35 cycles of 98°C for 15 second, 57°C for 15 second, 72°C (extension time according to the length of product) and ended with a final extension of 10 min at 72°C. Reaction condition of 678 bp segment amplification of TTSuV1 was 94°C for 5 min, followed by 40 cycles of 94°C for 30 s, 60°C for 45 sand 72°C for 3 min [20], with a final elongation step at 72°C for 7 min. Then products were cloned to T-Vector pMD19 (Simple) (TaKaRa) for sequencing.

Acknowledgments

We thank Xinglou Yang for instruction of mosquitoes processing, Zheng Zhu, Jun Wang and Kou Zheng for the help with 454 samples preparation and bioinformatics analysis. This work was supported by Major State Basic Research Development Program (2013FY113500), Important National Science & Technology Specific Projects (2012ZX10004219) and Wuhan Key Laboratory on Emerging Infectious Diseases and Biosafety.

Author Contributions

Conceived and designed the experiments: CS. Performed the experiments: CS. Analyzed the data: CS. Contributed reagents/materials/analysis tools: CS JX. Wrote the paper: CS YL XH BZ ZY.

References

1. Kuno G, Chang GJ, Tsuchiya KR, Karabatsos N, Cropp CB. Phylogeny of the genus *Flavivirus*. *Journal of virology*. 1998; 72(1):73–83. PMID: [9420202](#); PubMed Central PMCID: PMC109351.
2. Ho PS, Ng MM, Chu JJ. Establishment of one-step SYBR green-based real time-PCR assay for rapid detection and quantification of chikungunya virus infection. *Virology journal*. 2010; 7:13. doi: [10.1186/1743-422X-7-13](#) PMID: [20092632](#); PubMed Central PMCID: PMC2827404.
3. Forrester NL, Coffey LL, Weaver SC. Arboviral Bottlenecks and Challenges to Maintaining Diversity and Fitness during Mosquito Transmission. *Viruses*. 2014; 6(10):3991–4004. doi: [10.3390/v6103991](#) PMID: [25341663](#); PubMed Central PMCID: PMC4213574.
4. Molaei G, Andreadis TG, Armstrong PM, Anderson JF, Vossbrinck CR. Host feeding patterns of *Culex* mosquitoes and West Nile virus transmission, northeastern United States. *Emerg Infect Dis*. 2006; 12(3):468–74. doi: [10.3201/eid1205.051004](#) PMID: [16704786](#); PubMed Central PMCID: PMC3291451.
5. Ng TFF, Willner DL, Lim YW, Schmieder R, Chau B, Nilsson C, et al. Broad Surveys of DNA Viral Diversity Obtained through Viral Metagenomics of Mosquitoes. *Plos One*. 2011; 6(6). ARTN e20579 doi: [10.1371/journal.pone.0020579](#) PMID: [WOS:000291356400009](#).
6. Patel P, Landt O, Kaiser M, Faye O, Koppe T, Lass U, et al. Development of one-step quantitative reverse transcription PCR for the rapid detection of flaviviruses. *Virology journal*. 2013; 10:58. doi: [10.1186/1743-422X-10-58](#) PMID: [23410000](#); PubMed Central PMCID: PMC3616844.
7. Eshoo MW, Whitehouse CA, Zoll ST, Massire C, Pennella TT, Blyn LB, et al. Direct broad-range detection of alphaviruses in mosquito extracts. *Virology*. 2007; 368(2):286–95. doi: [10.1016/j.virol.2007.06.016](#) PMID: [17655905](#).
8. Delwart EL. Viral metagenomics. *Reviews in medical virology*. 2007; 17(2):115–31. doi: [10.1002/rmv.532](#) PMID: [17295196](#).

9. Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B. Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc Natl Acad Sci U S A*. 2005; 102(36):12891–6. doi: [10.1073/pnas.0504666102](https://doi.org/10.1073/pnas.0504666102) PMID: [16118271](https://pubmed.ncbi.nlm.nih.gov/16118271/); PubMed Central PMCID: PMC1200281.
10. Allander T, Jartti T, Gupta S, Niesters HG, Lehtinen P, Osterback R, et al. Human bocavirus and acute wheezing in children. *Clin Infect Dis*. 2007; 44(7):904–10. doi: [10.1086/512196](https://doi.org/10.1086/512196) PMID: [17342639](https://pubmed.ncbi.nlm.nih.gov/17342639/).
11. Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, Delwart EL. The fecal viral flora of wild rodents. *PLoS Pathog*. 2011; 7(9):e1002218. doi: [10.1371/journal.ppat.1002218](https://doi.org/10.1371/journal.ppat.1002218) PMID: [21909269](https://pubmed.ncbi.nlm.nih.gov/21909269/); PubMed Central PMCID: PMC3164639.
12. Finkbeiner SR, Allred AF, Tarr PI, Klein EJ, Kirkwood CD, Wang D. Metagenomic analysis of human diarrhea: viral detection and discovery. *PLoS Pathog*. 2008; 4(2):e1000011. doi: [10.1371/journal.ppat.1000011](https://doi.org/10.1371/journal.ppat.1000011) PMID: [18398449](https://pubmed.ncbi.nlm.nih.gov/18398449/); PubMed Central PMCID: PMC2290972.
13. Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, Delwart E. New DNA viruses identified in patients with acute viral infection syndrome. *Journal of virology*. 2005; 79(13):8230–6. doi: [10.1128/JVI.79.13.8230-8236.2005](https://doi.org/10.1128/JVI.79.13.8230-8236.2005) PMID: [15956568](https://pubmed.ncbi.nlm.nih.gov/15956568/); PubMed Central PMCID: PMC1143717.
14. Ng TF, Suedmeyer WK, Wheeler E, Gulland F, Breitbart M. Novel anellovirus discovered from a mortality event of captive California sea lions. *The Journal of general virology*. 2009; 90(Pt 5):1256–61. doi: [10.1099/vir.0.008987-0](https://doi.org/10.1099/vir.0.008987-0) PMID: [19264590](https://pubmed.ncbi.nlm.nih.gov/19264590/).
15. Coetzee B, Freeborough MJ, Maree HJ, Celton JM, Rees DJ, Burger JT. Deep sequencing analysis of viruses infecting grapevines: Virome of a vineyard. *Virology*. 2010; 400(2):157–63. doi: [10.1016/j.virol.2010.01.023](https://doi.org/10.1016/j.virol.2010.01.023) PMID: [20172578](https://pubmed.ncbi.nlm.nih.gov/20172578/).
16. Lopez-Bueno A, Tamames J, Velazquez D, Moya A, Quesada A, Alcamí A. High diversity of the viral community from an Antarctic lake. *Science*. 2009; 326(5954):858–61. doi: [10.1126/science.1179287](https://doi.org/10.1126/science.1179287) PMID: [19892985](https://pubmed.ncbi.nlm.nih.gov/19892985/).
17. Edwards RA, Rohwer F. Viral metagenomics. *Nature reviews Microbiology*. 2005; 3(6):504–10. doi: [10.1038/nrmicro1163](https://doi.org/10.1038/nrmicro1163) PMID: [15886693](https://pubmed.ncbi.nlm.nih.gov/15886693/).
18. Donaldson EF, Haskew AN, Gates JE, Huynh J, Moore CJ, Frieman MB. Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat. *Journal of virology*. 2010; 84(24):13004–18. doi: [10.1128/JVI.01255-10](https://doi.org/10.1128/JVI.01255-10) PMID: [20926577](https://pubmed.ncbi.nlm.nih.gov/20926577/); PubMed Central PMCID: PMC3004358.
19. Cortey M, Pileri E, Segales J, Kekkarainen T. Globalisation and global trade influence molecular viral population genetics of Torque Teno Sus Viruses 1 and 2 in pigs. *Veterinary microbiology*. 2012; 156(1–2):81–7. doi: [10.1016/j.vetmic.2011.10.026](https://doi.org/10.1016/j.vetmic.2011.10.026) PMID: [22101091](https://pubmed.ncbi.nlm.nih.gov/22101091/).
20. Cadar D, Kiss T, Adam D, Csagola A, Novosel D, Tuboly T. Phylogeny, spatio-temporal phylodynamics and evolutionary scenario of Torque teno sus virus 1 (TTSuV1) and 2 (TTSuV2) in wild boars: fast dispersal and high genetic diversity. *Veterinary microbiology*. 2013; 166(1–2):200–13. doi: [10.1016/j.vetmic.2013.06.010](https://doi.org/10.1016/j.vetmic.2013.06.010) PMID: [23850440](https://pubmed.ncbi.nlm.nih.gov/23850440/).
21. Liu JB, Guo LJ, Zhang L, Wei YW, Huang LP, Wu HL, et al. Three new emerging subgroups of torque teno sus viruses (TTSuVs) and co-infection of TTSuVs with porcine circovirus type 2 in China. *Virology journal*. 2013.
22. Zhai SL, Long JX, Wei WK, Chen QL, Luo ML, Lv DH, et al. High prevalence of torque teno sus virus in China and genetic diversity of the 5' non-coding region. *Arch Virol*. 2013; 158(7):1567–73. doi: [10.1007/s00705-013-1644-4](https://doi.org/10.1007/s00705-013-1644-4) PMID: [23435953](https://pubmed.ncbi.nlm.nih.gov/23435953/).
23. Cartwright SF, Lucas M, Huck RA. A small haemagglutinating porcine DNA virus. I. Isolation and properties. *Journal of comparative pathology*. 1969; 79(3):371–7. PMID: [4899939](https://pubmed.ncbi.nlm.nih.gov/4899939/).
24. Mayr A, Bachmann PA, Siegl G, Mahnel H, Sheffy BE. Characterization of a small porcine DNA virus. *Archiv fur die gesamte Virusforschung*. 1968; 25(1):38–51. PMID: [5729634](https://pubmed.ncbi.nlm.nih.gov/5729634/).
25. Streck AF, Bonatto SL, Homeier T, Souza CK, Goncalves KR, Gava D, et al. High rate of viral evolution in the capsid protein of porcine parvovirus. *The Journal of general virology*. 2011; 92(Pt 11):2628–36. doi: [10.1099/vir.0.033662-0](https://doi.org/10.1099/vir.0.033662-0) PMID: [21795474](https://pubmed.ncbi.nlm.nih.gov/21795474/).
26. Chen S, Cheng L, Zhang Q, Lin W, Lu X, Brannan J, et al. Genetic, biochemical, and structural characterization of a new densovirus isolated from a chronically infected *Aedes albopictus* C6/36 cell line. *Virology*. 2004; 318(1):123–33. doi: [10.1016/j.virol.2003.09.013](https://doi.org/10.1016/j.virol.2003.09.013) PMID: [14972541](https://pubmed.ncbi.nlm.nih.gov/14972541/).
27. Shike H, Dhar AK, Burns JC, Shimizu C, Jousset FX, Klimpel KR, et al. Infectious hypodermal and hematopoietic necrosis virus of shrimp is related to mosquito brevidensoviruses. *Virology*. 2000; 277(1):167–77. doi: [10.1006/viro.2000.0589](https://doi.org/10.1006/viro.2000.0589) PMID: [11062047](https://pubmed.ncbi.nlm.nih.gov/11062047/).
28. Afanasiev BN, Galyov EE, Buchatsky LP, Kozlov YV. Nucleotide sequence and genomic organization of *Aedes* densovirus. *Virology*. 1991; 185(1):323–36. PMID: [1833875](https://pubmed.ncbi.nlm.nih.gov/1833875/).

29. Boublik Y, Jousset FX, Bergoin M. Complete nucleotide sequence and genomic organization of the *Aedes albopictus* parvovirus (AaPV) pathogenic for *Aedes aegypti* larvae. *Virology*. 1994; 200(2):752–63. doi: [10.1006/viro.1994.1239](https://doi.org/10.1006/viro.1994.1239) PMID: [8178459](https://pubmed.ncbi.nlm.nih.gov/8178459/).
30. Zhai YG, Lv XJ, Sun XH, Fu SH, Gong ZD, Fen Y, et al. Isolation and characterization of the full coding sequence of a novel densovirus from the mosquito *Culex pipiens pallens*. *The Journal of general virology*. 2008; 89(Pt 1):195–9. doi: [10.1099/vir.0.83221-0](https://doi.org/10.1099/vir.0.83221-0) PMID: [18089743](https://pubmed.ncbi.nlm.nih.gov/18089743/).
31. O'Neill SL, Kittayapong P, Braig HR, Andreadis JTG, Gonzalez JP, Tesh I RB. Insect densoviruses may be widespread in mosquito cell lines. *The Journal of general virology*. 1995; 76:2067–74. PMID: [7636489](https://pubmed.ncbi.nlm.nih.gov/7636489/)
32. Jousset FX, Barreau C, Boublik Y, Cornet M. A parvo-like virus persistently infecting a C6/36 clone of *Aedes albopictus* mosquito cell line and pathogenic for *Aedes aegypti* larvae. *Virus Res*. 1993; 29(2):99–114. PMID: [8212862](https://pubmed.ncbi.nlm.nih.gov/8212862/).
33. Jousset FX, Baquerizo E, Bergoin M. A new densovirus isolated from the mosquito *Culex pipiens* (Diptera: culicidae). *Virus Res*. 2000; 67(1):11–6. PMID: [10773314](https://pubmed.ncbi.nlm.nih.gov/10773314/).
34. Ren X, Hoiczky E, Rasgon JL. Viral paratransgenesis in the malaria vector *Anopheles gambiae*. *PLoS Pathog*. 2008; 4(8):e1000135. doi: [10.1371/journal.ppat.1000135](https://doi.org/10.1371/journal.ppat.1000135) PMID: [18725926](https://pubmed.ncbi.nlm.nih.gov/18725926/); PubMed Central PMCID: PMC2500179.
35. Runckel C, Flenniken ML, Engel JC, Ruby JG, Ganem D, et al. Temporal Analysis of the Honey Bee Microbiome Reveals Four Novel Viruses and Seasonal Prevalence of Known Viruses, *Nosema* and *Crithidia*. *Plos One*. 2011; 6(6): e20656. doi: [10.1371/journal.pone.0020656.g001](https://doi.org/10.1371/journal.pone.0020656.g001) PMID: [21687739](https://pubmed.ncbi.nlm.nih.gov/21687739/)
36. Hogenhout SA, Redinbaugh MG, Ammar el D. Plant and animal rhabdovirus host range: a bug's view. *Trends in microbiology*. 2003; 11(6):264–71. PMID: [12823943](https://pubmed.ncbi.nlm.nih.gov/12823943/).
37. Mavale MS, Fulmali PV, Ghodke YS, Mishra AC, Kanojia P, Geevarghese G. Experimental transmission of Chandipura virus by *Phlebotomus argentipes* (diptera: psychodidae). *The American journal of tropical medicine and hygiene*. 2007; 76(2):307–9. PMID: [17297040](https://pubmed.ncbi.nlm.nih.gov/17297040/).
38. Mavale MS, Geevarghese G, Ghodke YS, Fulmali PV, Singh A, Mishra AC. Vertical and venereal transmission of Chandipura virus (Rhabdoviridae) by *Aedes aegypti* (Diptera: Culicidae). *J Med Entomol*. 2005; 42(5):909–11. PMID: [16363175](https://pubmed.ncbi.nlm.nih.gov/16363175/).
39. Kuwata R, Isawa H, Hoshino K, Tsuda Y, Yanase T, Sasaki T, et al. RNA splicing in a new rhabdovirus from *Culex* mosquitoes. *Journal of virology*. 2011; 85(13):6185–96. doi: [10.1128/JVI.00040-11](https://doi.org/10.1128/JVI.00040-11) PMID: [21507977](https://pubmed.ncbi.nlm.nih.gov/21507977/); PubMed Central PMCID: PMC3126488.
40. Nederbragt AJ. On the middle ground between open source and commercial software—the case of the Newbler program. *Genome biology*. 2014; 15(4):113. PMID: [25180324](https://pubmed.ncbi.nlm.nih.gov/25180324/); PubMed Central PMCID: PMC4054848.
41. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution*. 2013; 30(12):2725–9. doi: [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/); PubMed Central PMCID: PMC3840312.