



RESEARCH ARTICLE

Comparative Metagenomic Profiling of Viromes Associated with Four Common Mosquito Species in China

Han Xia¹ · Yujuan Wang^{1,2} · Chenyan Shi³ · Evans Atoni^{1,2} · Lu Zhao^{1,2} · Zhiming Yuan¹

Received: 14 November 2017 / Accepted: 12 January 2018
© Wuhan Institute of Virology, CAS and Springer Nature Singapore Pte Ltd. 2018

Abstract

Vast viruses are thought to be associated with mosquitoes. *Anopheles sinensis*, *Armigeres subalbatus*, *Culex quinquefasciatus*, and *Culex tritaeniorhynchus* are very common mosquito species in China, and whether the virome structure in each species is species-specific has not been evaluated. In this study, a total of 2222 mosquitoes were collected from the same geographic location, and RNAs were sequenced using the Illumina Miseq platform. After querying to the Refseq database, a total of 3,435,781, 2,223,509, 5,727,523, and 6,387,867 paired-end reads were classified under viral sequences from *An. sinensis*, *Ar. subalbatus*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus*, respectively, with the highest prevalence of virus-associated reads being observed in *Cx. quinquefasciatus*. The metagenomic comparison analysis showed that the virus-related reads were distributed across 26 virus families, together with an unclassified group of viruses. *Anelloviridae*, *Circoviridae*, *Genomoviridae*, *Iridoviridae*, *Mesoniviridae*, *Microviridae*, *Myoviridae*, *Parvoviridae*, *Phenuiviridae*, and *Podoviridae* were the top ten significantly different viral families among the four species. Further analysis reveals that the virome is species-specific in four mosquito samples, and several viral sequences which maybe belong to novel viruses are discovered for the first time in those mosquitoes. This investigation provides a basis for a comprehensive knowledge on the mosquito virome status in China.

Keywords Mosquito · Comparative metagenomics · Virome · China

Introduction

Mosquitoes, which can be classified into approximately 3500 species and further grouped into 41 genera, are the most common and important vectors for viruses.

Han Xia and Yujuan Wang have contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12250-018-0015-4>) contains supplementary material, which is available to authorized users.

✉ Zhiming Yuan
yzm@wh.iov.cn

- ¹ Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China
- ² University of Chinese Academy of Sciences, Beijing 100049, China
- ³ Department of Microbiology and Immunology, Laboratory of Viral Metagenomics, Rega Institute for Medical Research Leuven, KU Leuven - University of Leuven, 3000 Louvain, Belgium

Mosquitoes carry or transmit a wide variety of viruses which belong to two categories. The first category includes the viruses that can replicate in both mosquito and vertebrate cells, or can infect humans or other vertebrates. Most of the viruses in this group are known as pathogens of humans or animals, such as yellow fever virus (YFV), dengue virus (DENV), West Nile virus (WNV), chikungunya virus (CHIKV), Zika virus (ZIKV), Rift Valley virus (RVFV), *Akabane orthobunyavirus* (AKAV), Tembusu virus (TMUV), and Banna virus (BAV), within the families *Flaviviridae*, *Togaviridae*, *Reoviridae*, and *Phenuiviridae* (Conway et al. 2014). The second category comprises of the insect-specific viruses (ISVs), which only naturally infect mosquitoes, and replicate in mosquito cells in vitro, but cannot replicate in vertebrate cells, or infect humans or other vertebrates. They are widely distributed in the families of *Birnaviridae*, *Flaviviridae*, *Mesoniviridae*, *Parvoviridae*, *Peribunyaviridae*, *Reoviridae*, *Rhabdoviridae*, *Togaviridae*, and *Tymoviridae* (Bolling et al. 2015).

An. sinensis mainly inhabits rice fields, ditches, streams, irrigation canals, marshes, ponds, and ground

pools. They are a zoophilic mosquito which prefer to bite warm-blooded animals such as cattle (Zhu et al. 2013). *An. sinensis* is the main vector for malaria, but is also considered to be a vector for viruses such as Japanese encephalitis virus (JEV), BAV, and Kadipiro virus (KDV) (Liu et al. 2013; Zhang 1990). *Ar. subalbatus* commonly live in the places close to human dwellings, especially in sub-urban areas with poor sanitation and polluted water sites. They prefer to feed on humans or livestock (Rajavel 1992). *Ar. subalbatus* is known to be the vector of JEV and Getah virus (GETV) (Chen et al. 2000; Zheng et al. 2015). *Cx. quinquefasciatus* is widely distributed globally in populated regions and is often closely associated with bites on humans, other mammals, or birds (Farajollahi et al. 2011). *C. quinquefasciatus* is known to be an important vector of WNV. It also transmits St. Louis encephalitis virus, western equine encephalitis virus, and ZIKV (Samy et al. 2016). *Cx. tritaeniorhynchus* inhabits wells, ponds, and urban environments close to human settlements. It is known to be the main vector of JEV, and also a potential vector for BAV and TMUV (Liu et al. 2010; Self et al. 1973; Tang et al. 2013). All four mosquito species are extremely common and widespread in China.

Recently, the rapid development and application of next generation sequencing (NGS) technology has enabled broad surveys of viral diversity and has significantly increased our knowledge of viruses present in vectors, such as mosquitoes and ticks (Brinkmann et al. 2016; Xia et al. 2015). For example, detection of the presence of vertebrate viruses, originating from the blood of mosquitoes' host, which then accumulate in the mosquitoes' intestine during blood feeding, is possible (Chandler et al. 2015; Ng et al. 2011; Shi et al. 2015). In addition, recently, multiple unclassified mosquito-associated virus sequences, such as Wuhan mosquito virus, Xinzhou mosquito virus, Zhejiang mosquito virus, Zhee Mosquito virus, Wutai Mosquito virus, and *Cx. tritaeniorhynchus* rhabdovirus, have been found (Shi et al. 2016). In our previous study, a metagenomic survey of mosquitoes collected in different cities in the Hubei province, the results indicated a high abundance and diversity of viruses in mosquitoes, and several novel viruses or sequences were identified (Shi et al. 2015). However, the characteristic features of the virome associated with different mosquito species in same geographic locations has not been well investigated and evaluated.

In this study, a metagenomic comparison analysis for the virome structure associated with four different mosquito species (*An. sinensis*, *Ar. subalbatus*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus*) collected in Yichang region was carried out, and the results indicated distinct virome profiles in each species, indicating that the differences among viromes are species-specific.

Materials and Methods

Mosquito Sampling and Taxonomy Identification

A total of 2222 adult mosquitoes were collected from sites adjacent to human settlements, cowsheds, and pigpens, in Yichang region of Hubei Province in China. Sampling was done in the months of June to September 2014. No protected species were sampled. The mosquito capture method involved the use of electrical mosquito aspirators (Yalin, China). The collected mosquitoes were transported to laboratory alive inside net traps, then euthanized by freezing at -20°C for 30 min. They were then morphologically classified and further grouped into pools of approximately 50 adult mosquitoes per pool. The mosquito pools were subsequently washed once in 70% ethanol and distilled water, for 5 min each, to remove environmental contaminants.

Nucleic Acid Extraction and Sequencing

The pools of mosquitoes in each species were homogenized and centrifuged as described previously (Shi et al. 2015). After centrifugation, the supernatant was filtered using a $0.22\text{-}\mu\text{m}$ polycarbonate filter membrane (Millipore, Billerica, USA) to remove bacteria and other debris present. RNA was extracted from $140\ \mu\text{L}$ of supernatant. The QIAamp viral RNA extraction kit (Qiagen, Hilden, Germany) was used for RNA extraction according to the manufacturer's instructions. The final RNA elution was done with $60\ \mu\text{L}$ AVE buffer. The RNA samples were sent to the Shanghai Biotechnology Corporation for library construction and sequencing by paired-end mode through the Illumina Miseq platform (Illumina, San Diego, USA).

Data Analysis

Sequencing adaptors were trimmed and reads with low quality or match to the mosquito genome were discarded, to acquire clean reads for each sample. De novo assembly was then used to obtain contigs. These steps were conducted by Shanghai Biotechnology Corporation.

The Kaiju web server was used for taxonomic classification of the reads and contigs, and the Greedy mode was chosen with SEG filter, minimum match length = 11, minimum match score = 75, and allowed mismatches = 5 (Menzel et al. 2016). The Kaiju RefSeq Genome (protein sequences from 7065 complete bacterial and archaeal genomes, and 9334 viral genomes from NCBI RefSeq, updated 2017-05-16) was selected as the reference database (<http://kaiju.binf.ku.dk/>). Blastx (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to make a double

confirmation for the contigs matching to viral sequences. The extraction for the information of assigned reads, which classified them to a viral family and species, together with other statistical analyses, were conducted in R, version 3.2.5 (<https://www.r-project.org/>). Metagenomic comparative analysis was conducted through MetaComp (<http://cqb.pku.edu.cn/ZhuLab/MetaComp/>) under the multi-sample statistic method (Zhai et al. 2017).

Results

Taxonomic Identification of Mosquitoes

A total of 2222 mosquitoes were collected and identified to species level through morphological identification under a stereomicroscope. They were classified into four species: *An. sinensis*, *Ar. subalbatus*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus*, with 387, 852, 745, and 238 mosquitoes from each species, respectively. The four species are very common in the Yichang region, Hubei, and the dominant species in 2014 were *Ar. subalbatus* (38.3%) followed by *Cx. quinquefasciatus* (33.5%).

Illumina Sequencing Results and Viral-Related Reads/Contigs in Four Mosquito Species

A total of 6,535,325, 7,183,603, 6,714,707 and 8,909,955 clean, paired reads and 251, 198, 284, and 186 assembled contigs were generated from *An. sinensis*, *Ar. subalbatus*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus*, respectively, through Illumina Miseq sequencing. After querying of the RefSeq Complete Genomes database through the Kaiju Web Server, the result indicated that there were many reads that corresponded to known viral sequences (Table 1). The percentage of the reads that matched to viral sequences were 52.5%, 30.9%, 85.3%, and 71.1% in *An. sinensis*, *Ar. subalbatus*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus* respectively, and highest prevalence of virus associated reads were observed in *Cx. quinquefasciatus*.

Comparative Analysis for the Viral Family Profile Based on Reads

Taxonomic classification was conducted based on viral families. The results indicate that the virus-related reads present in the four mosquito viromes are distributed across 26 virus families, together with an unclassified group (Supplementary Table S1).

Anelloviridae, *Circoviridae*, *Genomoviridae*, *Iridoviridae*, *Mesoniviridae*, *Microviridae*, *Myoviridae*, *Parvoviridae*, *Phenuiviridae*, and *Podoviridae* were the top ten significantly different viral families (Fig. 1). According to the detailed analysis results in Supplementary Table S1, many reads related to *Iridoviridae*, *Poxviridae*, *Totiviridae*, and *Peribunyaviridae* were found in *An. sinensis*, while they were rarely detected in the other three species. In addition, the reads that matched to *Phenuiviridae* and *Siphoviridae* were identified in *Ar. subalbatus*, however rarely found in the other species.

To detect differences in virome structures among the different mosquito species, a taxonomic heatmap and hierarchical cluster analysis were conducted based on a distance matrix (Fig. 2). The heatmap analysis based on virus family showed that each species had a special virome structure. For instance, in *An. sinensis* mosquitoes, *Parvoviridae*, *Reoviridae*, *Poxviridae*, *Iridoviridae*, *Anelloviridae*, *Totiviridae*, *Myoviridae*, *Inoviridae*, and *Peribunyaviridae* were found to be present in abundant quantities while in *Ar. subalbatus*, *Iflaviridae*, *Microviridae*, *Siphoviridae*, *Circoviridae*, and *Phenuiviridae* were the most abundant virus families. However, the abundant viral families in *Cx. quinquefasciatus* species were found to be *Genomoviridae*, *Ascoviridae*, *Tymoviridae*, *Marseilleviridae*, and *Polydnaviridae*. In *Cx. tritaeniorhynchus* mosquitoes, *Mesoniviridae* and *Geminiviridae* were detected in abundant quantities. The cluster results showed *Cx. quinquefasciatus* and *Cx. tritaeniorhynchus* grouped together, which indicated the correlation of the virome structures in these two species were closer than they were to the other species, to some extent.

Table 1 Reads/contigs in each mosquito species.

Species	Clean paired reads	Classified to viruses (viral reads/clean reads)	Contigs (average length > 500 bp)	Classified to viruses (viral contigs/contigs)
<i>Anopheles sinensis</i>	6,535,325	3,435,781 (52.6%)	251	33 (13.1%)
<i>Armigeres subalbatus</i>	7,183,603	2,223,509 (30.9%)	198	58 (29.3%)
<i>Culex quinquefasciatus</i>	6,714,707	5,727,523 (85.3%)	284	197 (69.4%)
<i>Culex tritaeniorhynchus</i>	8,909,955	6,387,867 (71.7%)	186	107 (57.5%)

Fig. 1 Top ten significantly different viral families among four mosquito species. Frequency is the rate of the number of reads belong to a viral family to the total number of viral reads in one mosquito sample. The frequencies of *Anelloviridae* and *Parvoviridae* in four samples dramatically fluctuated.

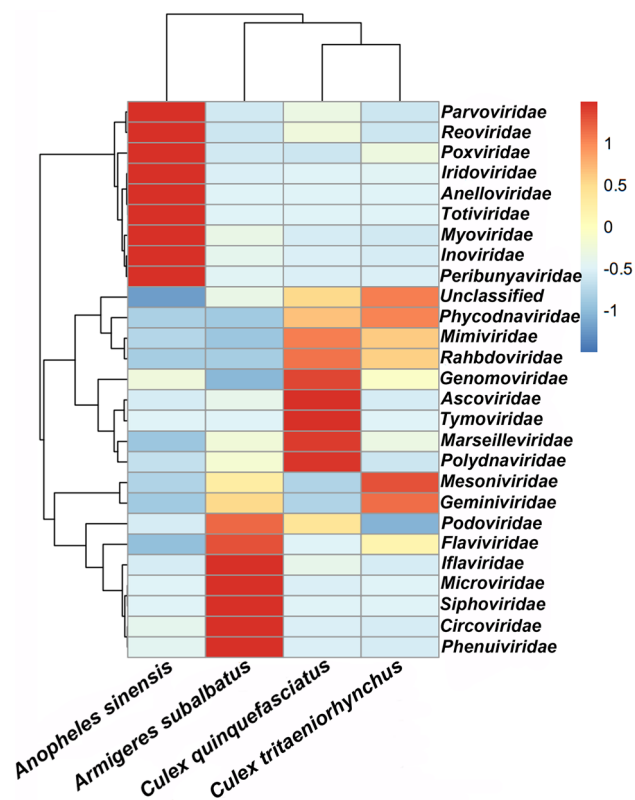
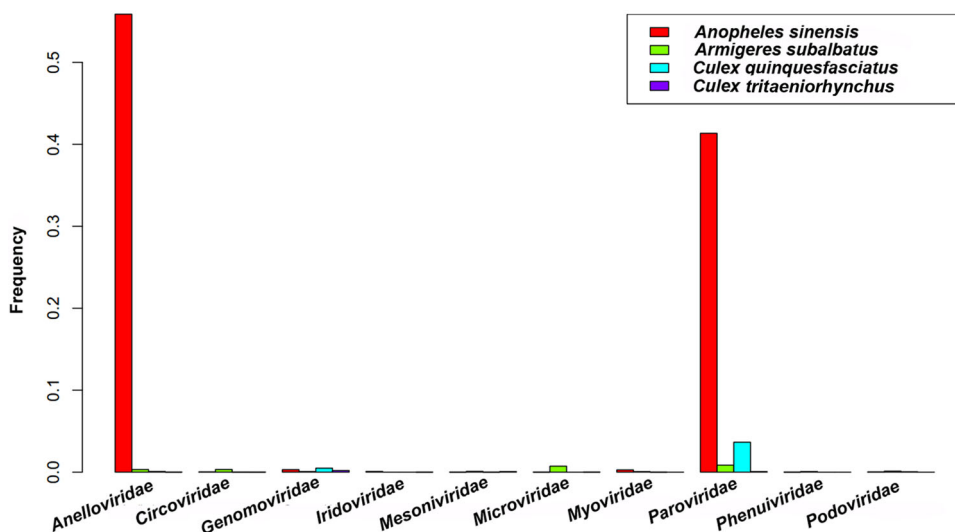


Fig. 2 Heatmap based on 26 viral families. Heatmap was constructed based on the distance matrix, which was calculated by Euclidean distance method according to the number of reads belonging to each viral family in four mosquito species. The X axis represents the sample names of mosquito species, the Y axis represents the names of viral families. The color red to blue stands for the highest to lowest abundant of viral reads in each species according to viral family.

For example, the *Phycodnaviridae*, *Mimiviridae* and *Rhabdoviridae* viral families were found to be in abundant or moderate levels in both *Culex* species.

The Viral Species Profile in Four Mosquito Species Based on Assembled Contigs

The information for contigs that matched to the known viruses in each sample is summarized in Table 2. In *An. sinensis*, three types of Torque teno sus virus (1a, 1b, and k2) were detected, however, none of them were detected in *Cx. tritaeniorhynchus*. Three types of mesonivirus were found in *Cx. tritaeniorhynchus*, and two types in *Ar. subalbatus*, but none of them were detected in *An. sinensis* or *Cx. quinquefasciatus*. The Wuhan fly virus was only detected in *An. sinensis*. BAV was found in *An. sinensis* and *Cx. tritaeniorhynchus*. Large amount of contigs matched to *Cx. tritaeniorhynchus* rhabdovirus, Wenzhou sobemo-like virus 3, and Wuhan mosquito virus 8 in the two *Culex* species. Hubei mosquito virus 2 was found in all four species. Many newly identified viruses, such as the Hubei mosquito virus, Hubei tombus-like virus, Wenzhou sobemo-like, and Zhejiang mosquito virus reported by Shi et al. (2016), were detected in our mosquito samples (Table 2).

In addition, several viral sequences which maybe belong to novel viruses (low identity to known virus) for the first time presented in mosquitoes were discovered (Table 2). For example, the Tanay virus was only previously reported in *Culex* spp. or *Armigeres* spp. in the Philippines (Li et al. 2014). However, there was one contig matched with Tanay virus (with 46% identity) in our *An. sinensis* sample. Several contigs hit to Wuhan fly virus 1, Wuhan ant virus, Beihai mantis shrimp virus 6, Hubei picorna-like virus 41, Hubei rhabdo-like virus 1, Sanxia water strider virus 8, or Wuhan house centipede virus 6, with the identity ranging from 29% to 76%, which was the first time that these viruses have been found in mosquitoes. In future work, we will try to isolate them and conduct further analysis.

Table 2 Information on contig hits to the virus database in four mosquito species.

Family	Host	Species	No. of contigs match to virus (Blastx e-value < 0.00001)			
			<i>Anopheles sinensis</i> (identity)	<i>Armigeres subalbatus</i> (identity)	<i>Culex quinquefasciatus</i> (identity)	<i>Culex tritaeniorhynchus</i> (identity)
<i>Anelloviridae</i>	Vertebrates, human	Torque teno sus virus 1a	3 (72%–100%)	2 (79%–83%)	3 (79%–97%)	0
		Torque teno sus virus 1b	4 (97%–100%)	0	0	0
		Torque teno sus virus k2	1 (45%)	0	0	0
<i>Circoviridae</i>	Vertebrates	Canary circovirus	0	1 (40%)	0	0
<i>Genomoviridae</i>	NA	<i>Poecile atricapillus</i> GI tract-associated gemycircularvirus	0	0	2 (80%–83%)	1 (87%)
		Dragonfly associated gemycircularvirus 1	0	0	0	1 (47%)
		Soybean leaf-associated gemycircularvirus 1	0	0	1 (68%)	0
	Plant	Common bean-associated gemycircularvirus	1 (61%)	0	2 (48%–50%)	0
<i>Mesoniviridae</i>	Invertebrates	Alphamesonivirus 1	0	1 (51%)	0	
		Alphamesonivirus 3	0	0	0	1 (67%)
		Alphamesonivirus 4	0	1 (50%)		3 (54%)
<i>Microviridae</i>	Bacteria	Microviridae Fen7918_21	0	1 (48%)	0	0
		Marine gokushovirus	0	1 (52%)	0	0
		Gokushovirinae GNX3R	0	1 (49%)	0	0
		Gokushovirinae Bog1183_53	0	1 (47%)	0	0
<i>Myoviridae</i>	Bacteria	Enterobacteria phage phi92	1 (81%)	0	0	0
		Escherichia virus CC31	1 (69%)	0	0	0
<i>Parvoviridae</i>	Invertebrates	<i>Aedes albopictus</i> densovirus	0	0	0	1 (87%)
		<i>Anopheles gambiae</i> densovirus	0	1 (89%)	0	0
		Mosquito densovirus BR/07	0	1 (90%)	0	1
		<i>Planococcus citri</i> densovirus	1 (39%)	1 (51%)	0	0
	Vertebrates	Porcine partetravirus	0	2 (99%–100%)	0	0
		Porcine parvovirus 2	6 (83%–100%)	4 (98%–100%)	4 (99%)	0
<i>Peribunyaviridae</i>	Invertebrates	Wuhan fly virus 1	1 (29%)	0	0	0
<i>Podoviridae</i>	Bacteria	Phage vB_EcoP_SU10	0	1 (72%)	0	0
		Erwinia phage PEp14	0	0	1 (78%)	0
<i>Reoviridae</i>	vertebrates invertebrates human	Banna virus	7 (63%–91%)	0	3 (58%–67%)	0
<i>Rhabdoviridae</i>	Invertebrates	<i>Culex tritaeniorhynchus</i> rhabdovirus	0	0	55 (94%–100%)	46 (97%–100%)
		Wuhan ant virus	0	2 (38%–76%)	0	0
		Wuhan Mosquito Virus 9	0	0	17 (88%–100%)	2 (100%)
<i>Siphoviridae</i>	Bacteria	Escherichia virus K1-dep (4)	0	1 (70%)	0	0
		<i>Salmonella</i> virus SP31	0	1 (75%)	0	0
<i>Totiviridae</i>	Protozoa	Red clover powdery mildew-associated totivirus 2	1 (38%)	0	0	0
Unclassified	Invertebrates	Beihai mantis shrimp virus 6	2 (53%–55%)	0	0	0
		Hubei mosquito virus 2	3 (75%–100%)	3 (86%–97%)	26 (73%–100%)	4 (70%–100%)
		Hubei mosquito virus 4	0	0	14 (98%–100%)	0
		Hubei picorna-like virus 41	0	2 (46%–60%)	0	0
		Hubei rhabdo-like virus 1	0	1 (37%)	0	0
		Hubei sobemo-like virus 41	0	2 (99%–100%)	0	0
		Hubei tombus-like virus 20	0	11 (97%–100%)	0	0

Table 2 continued

Family	Host	Species	No. of contigs match to virus (Blastx e-value < 0.00001)			
			<i>Anopheles sinensis</i> (identity)	<i>Armigeres subalbatus</i> (identity)	<i>Culex quinquefasciatus</i> (identity)	<i>Culex tritaeniorhynchus</i> (identity)
		Hubei tombus-like virus 28	0	1 (54%)	0	1 (54%)
		Hubei toti-like virus 10	0	11 (99%–100%)	0	0
		Sanxia water strider virus 8	0	1 (51%)	0	0
		Shahe heteroptera virus 3	0	1 (29%)	0	0
		Tanay virus	1 (46%)	0	0	0
		Wenzhou sobemo-like virus 3	0	0	14 (70%–99%)	25 (78%–100%)
		Wenzhou sobemo-like virus 4	0	1 (77%)	1 (63%)	1 (83%)
		Wenzhou tombus-like virus 11	0	0	5 (99%–100%)	8 (89%–100%)
		Wuhan house centipede virus 6	0	0	1 (50%)	0
		Wuhan mosquito virus 2	0	0	0	1 (100%)
		Wuhan mosquito virus 8	0	0	42 (96%–100%)	10 (97%–100%)
		Zhejiang mosquito virus 1	0	0	4 (80%–84%)	0
Fungi		<i>Sclerotinia sclerotiorum</i> hypovirulence associated DNA virus 1	0	1 (73%)	0	0

Discussion

Mosquitoes are known to obtain blood meal from vertebrates, which in turn are known to be natural reservoirs of the majority of viruses that are transmitted by invertebrate vectors. Hence, the diversity for mosquito virome could be influenced by the communities of vertebrates present. Most of the reported metagenomic analyses of the mosquito virome compared the diversity among the mosquitoes from different locations or ecological systems (Chandler et al. 2015; Frey et al. 2016; Ng et al. 2011; Shi et al. 2015), in which it is difficult to infer whether the differences between viromes were caused by the ecological communities in different locations or by the mosquito species. Here, we describe the biodiversity of the viral communities associated with *An. sinensis*, *Ar. subalbatus*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus* mosquitoes, which inhabit Yichang and commonly attack people or livestock in this region. We detected the virome differences among the four species and the results indicated that each species had a specific taxonomic structure, and that they contained sequences related to a broad range of animal, plant, insect, bacterial, fungal, and protozoan viruses. In addition, the clustering algorithms demonstrated that the distance between the viromes of *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus*, were much closer than between the other species, which may be because of them being in the same genus.

BAV is the member of genus *Seadornavirus* in Family *Reoviridae*, and its genome has 12 dsRNA segments. It was first isolated from the cerebrospinal fluid and serum of encephalitis patients in Xishuangbanna, Yunnan, in 1987. It is considered to be a new pathogen, causing human viral encephalitis and fever (Liu et al. 2010; Xu et al. 1990). Subsequently, BAV has been

isolated from 10 mosquito species in 3 genera (*Cx. tritaeniorhynchus*, *Cx. pipiens pallens*, *Cx. annulus*, *Cx. pseudovishnui*, *Cx. modestus*, *An. sinensis*, *Ae. vagus*, *Ae. albopictus*, *Ae. vexans*, and *Ae. dorsalis*) mosquitoes from China (Liu et al. 2010). In this study, the virus sequences associated to BAV within the family *Reoviridae* were detected in all four mosquito species tested, and the highest abundance was found in the *Ar. subalbatus* mosquito species (Supplementary Table S1 and Table 2), which provides further evidence for the widespread nature of BAV in different mosquito species. In addition, this result indicated the potential risks of BAV infection in the local residents of Yichang and that serological study should be conducted in the future.

Mesoniviridae is a newly classified viral family. The members in this family consist of ssRNA of ~ 20 kb genome size, and they replicate in mosquitoes. Mesoniviruses have been proposed to provide an evolutionary link between nidovirus lineages, with small and large RNA genome sizes (Lauber et al. 2012; Nga et al. 2011). A novel species of mesonivirus named Yichang virus was isolated from *Culex* spp. from the sample which was the same batch used for the NGS analysis here (Wang et al. 2017). The metagenomic study showed the *Mesoniviridae* related sequences were present in all four mosquito species, and they had a much higher prevalence in *Ar. subalbatus* and *Cx. tritaeniorhynchus*, with identity ranging from 50% to 67% (Supplementary Table S1 and Table 2), which indicated that the mesonivirus is commonly distributed in mosquitoes in this region.

Viral sequences belonging to unclassified virus were detected in all four species, but the prevalence of

unclassified virus was much higher in the sample from *Ar. subalbatius*, *Cx. quinquefasciatus* and *Cx. tritaeniorhynchus* when compared to *An. sinensis*. Most of these sequences are closely related to the newly identified virus by Shi et al. (2016), such as the Wuhan mosquito virus, Hubei mosquito virus, Zhejiang mosquito virus, and Wenzhou sobemo-like virus. This indicated that these types of viruses are extremely common in mosquitoes, and further work for towards their isolation and characterization should be done to investigate the role of their presence in mosquito.

In our previous study, the presence of vertebrate viruses, such as *Anelloviridae* and *Parvoviridae* originating from the blood of the mosquitoes' host, were found in the mosquito virome. A large number of viral sequences, which were closely related to the torque teno sus virus (distributed in pigs) or porcine parvovirus, were found in the sample from *Anopheles* spp. This might be because of pigs being one of the main hosts for the blood feeding of *Anopheles* mosquito, and that the infection rate for these two types of viruses in pigs is high in this region.

In summary, we sequenced and analyzed the mosquito virome from four species, and we found that the virome structure varied greatly among mosquito species (*An. sinensis*, *Ar. subalbatius*, *Cx. quinquefasciatus*, and *Cx. tritaeniorhynchus*) in the same geographic regions. This approach could help us to gain a better understanding of the virus flora in mosquito species.

Acknowledgements We thank the Hubei Provincial Center for Disease Control and Prevention for assistance in mosquito sampling within the Hubei Province. This work was supported by the Ministry of Science and Technology of China (Science and Technology Basic Work Program 2013FY113500).

Author Contributions HX and ZY designed the experiments. YW, and LZ, carried out the experiments. HX, YW, CS, and AE analyzed the data. HX, YW and ZY wrote the paper. HX and CS checked and finalized the manuscript. All authors read and approved the final manuscript.

Compliance with Ethics Standards

Conflict of interest The authors declare that they have no conflict of interest.

Animal and Human Rights Statement This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Bolling B, Weaver S, Tesh R, Vasilakis N (2015) Insect-specific virus discovery: significance for the arbovirus community. *Viruses* 7:4911–4928. <https://doi.org/10.3390/v7092851>

- Brinkmann A, Nitsche A, Kohl C (2016) Viral metagenomics on blood-feeding arthropods as a tool for human disease surveillance. *Int J Mol Sci.* <https://doi.org/10.3390/ijms17101743>
- Chandler JA, Liu RM, Bennett SN (2015) RNA shotgun metagenomic sequencing of Northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Front Microbiol* 6:185. <https://doi.org/10.3389/fmicb.2015.00185>
- Chen WJ, Dong CF, Chiou LY, Chuang WL (2000) Potential role of *Armigeres subalbatius* (Diptera: Culicidae) in the transmission of Japanese encephalitis virus in the absence of rice culture on Liuchiu islet, Taiwan. *J Med Entomol* 37:108–113. <https://doi.org/10.1603/0022-2585-37.1.108>
- Conway MJ, Colpitts TM, Fikrig E (2014) Role of the vector in arbovirus transmission. *Annu Rev Virol* 1:71–88. <https://doi.org/10.1146/annurev-virology-031413-085513>
- Farajollahi A, Fonseca DM, Kramer LD, Marm Kilpatrick A (2011) “Bird biting” mosquitoes and human disease: a review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infect Genet Evol* 11:1577–1585. <https://doi.org/10.1016/j.mecgid.2011.08.013>
- Frey KG, Biser T, Hamilton T, Santos CJ, Pimentel G, Mokashi VP, Bishop-Lilly KA (2016) Bioinformatic characterization of mosquito viromes within the Eastern United States and Puerto Rico: discovery of novel viruses. *Evol Bioinform* 12:1–12. <https://doi.org/10.4137/EBO.S38518>
- 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. <https://doi.org/10.1007/s00705-012-1295-x>
- Li W, Cao Y, Fu S, Wang J, Li M, Jiang S, Wang X, Xing S, Feng L, Wang Z, Shi Y, Zhao S, Wang H, Wang Z, Liang G (2014) Tahyna virus infection, a neglected arboviral disease in the Qinghai-Tibet Plateau of China. *Vector Borne Zoonotic Dis* 14:353–357. <https://doi.org/10.1089/vbz.2013.1351>
- Liu H, Li MH, Zhai YG, Meng WS, Sun XH, Cao YX, Fu SH, Wang HY, Xu LH, Tang Q, Liang GD (2010) Banna virus, China, 1987–2007. *Emerg Infect Dis* 16:514–517. <https://doi.org/10.3201/eid1603.091160>
- Liu H, Lu HJ, Liu ZJ, Jing J, Ren JQ, Liu YY, Lu F, Jin NY (2013) Japanese encephalitis virus in mosquitoes and swine in Yunnan province, China 2009–2010. *Vector Borne Zoonotic Dis* 13:41–49. <https://doi.org/10.1089/vbz.2012.1016>
- Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>
- Ng TFF, Willner DL, Lim YW, Schmieder R, Chau B, Nilsson C, Anthony S, Ruan Y, Rohwer F, Breitbart M (2011) Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS ONE* 6:e20579. <https://doi.org/10.1371/journal.pone.0020579>
- Nga PT, de Parquet MC, 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. <https://doi.org/10.1371/journal.ppat.1002215>
- Rajavel AR (1992) Larval habitat of *Armigeres subalbatius* (COQ) and its characteristics in Pondicherry. *Southeast Asian J Trop Med Public Health* 23:470–473
- Samy AM, Elaagip AH, Kenawy MA, Ayres CFJ, Peterson AT, Soliman DE (2016) Climate change influences on the global potential distribution of the mosquito *Culex quinquefasciatus*, vector of West Nile virus and lymphatic filariasis. *PLoS ONE* 11:e0163863. <https://doi.org/10.1371/journal.pone.0163863>

- Self LS, Shin HK, Kim KH, Lee KW, Chow CY, Hong HK (1973) Ecological studies on *Culex tritaeniorhynchus* as a vector of Japanese encephalitis. *Bull World Health Organ* 49:41–47
- 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:e0129845. <https://doi.org/10.1371/journal.pone.0129845>
- Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ (2016) Redefining the invertebrate RNA virosphere. *Nature* 540:539–543
- Tang Y, Diao Y, Chen H, Ou Q, Liu X, Gao X, Yu C, Wang L (2013) Isolation and genetic characterization of a tembusu virus strain isolated from mosquitoes in Shandong, China. *Transbound Emerg Dis* 62:209–216. <https://doi.org/10.1111/tbed.12111>
- Wang Y, Xia H, Zhang B, Liu X, Yuan Z (2017) Isolation and characterization of a novel mesonivirus from *Culex* mosquitoes in China. *Virus Res* 240:130–139. <https://doi.org/10.1016/j.virusres.2017.08.001>
- Xia H, Hu C, Zhang D, Tang S, Zhang Z, Kou Z, Fan Z, Bente D, Zeng C, Li T (2015) Metagenomic profile of the viral communities in *Rhipicephalus* spp. ticks from Yunnan, China. *PLoS ONE* 10:e0121609. <https://doi.org/10.1371/journal.pone.0121609>
- Xu P, Wang Y, Zuo J, Lin J, Xu P (1990) New orbiviruses isolated from patients with unknown fever and encephalitis in Yunnan province. *Chin J Virol* 6:27–33
- Zhai P, Yang L, Guo X, Wang Z, Guo J, Wang X, Zhu H (2017) MetaComp: comprehensive analysis software for comparative meta-omics including comparative metagenomics. *BMC Bioinformatics* 18:434. <https://doi.org/10.1186/s12859-017-1849-8>
- Zhang HL (1990) The natural infection rate of mosquitoes by Japanese encephalitis B virus in Yunnan province. *Zhonghua Yu Fang Yi Xue Za Zhi* 24:265–267 (in Chinese)
- Zheng Y, Cao Y, Fu S, Cheng J, Zhao J, Dai P, Kong X, Liang G (2015) Isolation and identification of mosquito-borne arboviruses in Yuncheng city, Shanxi province, 2012. *Zhonghua Liu Xing Bing Xue Za Zhi* 36:368–373 (in Chinese)
- Zhu G, Xia H, Zhou H, Li J, Lu F, Liu Y, Cao J, Gao Q, Sattabongkot J (2013) Susceptibility of *Anopheles sinensis* to *Plasmodium vivax* in malarial outbreak areas of central China. *Parasit Vectors* 6:176. <https://doi.org/10.1186/1756-3305-6-176>