Defining Virus-Carrier Networks That Shape the Composition of the Mosquito Core Virome of an Ecosystem

Kostantinos Konstantinidis
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Nikolas Dovrolis
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Adamantia Kouvela
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Katerina Kassela
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Maria Goreti Rosa Freitas
FIOCRUZ: Fundacao Oswaldo Cruz
Andreas Nearchou
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Michael de Courcy Williams
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Stavroula Veletza
Democritus University of Thrace: Demokriteio Panepistemio Thrakes
Ioannis Karakasiliotis (ioakarak@med.duth.gr)
Democritus University of Thrace: Demokriteio Panepistemio Thrakes https://orcid.org/0000-0003-3004-8104

Research

Keywords: Mosquitoes, virome, ecosystem, metagenomics, vectors

Posted Date: February 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-229254/v1

Abstract

Background

Mosquitoes are the most important vectors of emerging infectious diseases. During the past decade, our understanding of the diversity of viruses they carry has greatly expanded. Most of these viruses are considered mosquito-specific, while there is increasing evidence that these viruses may affect mosquito vector potential. Metagenomics approaches have focused on specific mosquito species for the identification of what is called core virome. However, in most ecosystems, multiple species may participate in virus emergence and circulation, while there is lack of understanding on the viruses-carrier/host network for both vector-borne and mosquito-specific viruses.

Results

Here, we studied the core mosquito virome in a diverse ecosystem comprised of 24 different mosquito species. Analysis of these 24 diverse viromes resulted in the discovery of 35 viruses with known genetic traits and 9 novel viruses. Comparison of the viromes of the 24 individual species revealed novel relationships between mosquito species and virus families, as most of the mosquito species had never been analysed in the past. Groups of related viruses and mosquito species from multiple genera formed a complex network in the ecosystem. Analyses of whole traps of mosquitoes of variable composition not only showed a stable core virome for each species but also a relationship between mosquito population and virome composition.

Conclusions

Our study highlighted the importance of a holistic approach regarding mosquito viromes in rich and diverse ecosystems. Our data further supported the idea of a stable core virome, characteristic of each mosquito species. The remarkable stability of the core virome seemed to determine the composition of the total mosquito core virome of a habitat in the ecosystem.

Background

Arboviruses pose a significant threat to public health worldwide causing epidemics such as Dengue, West Nile, Zika & chikungunya diseases with considerable repercussions for the infected hosts. Undoubtedly, tropical, rural and agricultural areas are plagued by arboviruses, while climate change, urbanization and globalization affect the geographic and seasonal patterns of vector-borne diseases, facilitating efficient reproduction of insects and the pathogens they may carry. Mosquitoes are major vectors for most known vector-borne diseases. Global surveillance data have reported a considerable increase of mosquito-borne diseases and pest-control of mosquitoes is a top priority during their activity period to prevent future outbreaks. In parallel to important human and other animal pathogens, mosquitoes carry a significant number of mosquito-specific viruses that either are limited to the insect host or no other secondary host has been identified to date. Viruses, that are considered to be insect-specific, are usually transmitted

vertically, through generations, and are recently of great interest as modulators of mosquito vector competency. Nhumirim virus, Eilat virus, Palm Creek virus, cell fusing agent virus and Phasi Charoen-like virus have been shown to affect West Nile, Zika and dengue virus replication *in vitro* and *in vivo* (Goenaga et al., 2015; Hall-Mendelin et al., 2016; Kenney et al., 2014; Nasar et al., 2015; Schultz et al., 2018).

Significant efforts have been made to elucidate the virome of insects in endeavors to characterize the constituents of the "virosphere" of natural habitats and identify sources of potential emerging infectious agents. Two key publications detailed the virosphere of several provinces in China. Li and colleagues analyzed 70 species from four different arthropod classes (*Insecta, Arachnida, Chilopoda, Malacostraca*) through RNA sequencing, discovering 112 novel viruses within Hubei, Zhejiang, Beijing and Xinjiang provinces of China (Li et al., 2015). Shi and colleagues profiled the transcriptomes of over 220 invertebrate species sampled across nine animal phyla and reported the discovery of 1,445 RNA viruses within the same provinces of China, filling virus phylogeny gaps (Shi et al., 2016). As mosquitoes are a top priority, considering their vector potential, several studies have focused on metagenomics approaches to define the mosquito virome in pools of mosquitoes that contain either multiple or single species. Mosquito species, usually belonging to *Aedes (Ae.)*, *Anopheles (An.)* and *Culex (Cx.)* genera that serve as important vectors worldwide have been analyzed in the past for the presence of viruses. Viruses belonging to the families of *Flaviviridae, Togaviridae, Bunyaviridae, Rhabdoviridae, Mesoniviridae, Tymoviridae, Birnaviridae, Reoviridae, Nodaviridae, Iridoviridae, Permutotetraviridae, Phenuiviridae, Anelloviridae, Arteriviridae, Peribunyaviridae, Parvoviridae, Iflaviridae, Orthomyxoviridae, Totiviridae, Chrysoviridae, Picornaviridae, Narnaviridae* and several unclassified others have been identified (Agboli et al., 2019; Atoni et al., 2020; Shi et al., 2019; Shi et al., 2017). The vast majority of the identified viruses are RNA viruses that encompass segmented or non-segmented RNA genomes. Comparative analysis of different mosquito species in the same habitat has highlighted a virome with a distinct core composition (Shi et al., 2019). Mosquito species that belong to the same genus present less distinctive core viromes (Shi et al., 2017), although in depth comparison has identified virome patterns that differentiate related species (Pettersson et al., 2019).

Comparison of core viromes among mosquito species in a habitat are a first step in defining the ecological dynamics of mosquito-associated viruses that may shape the ecology and transmission of pathogens, particularly of human or animal health importance such as West-Nile virus (Hall-Mendelin et al., 2016; Hobson-Peters et al., 2013; Vasilakis and Tesh, 2015). Such comparisons are usually restricted to between two or just a few mosquito species within a particular ecosystem, target normally at the medically important species of the region (Fauver et al., 2016; Pettersson et al., 2019; Shi et al., 2019). A meta-analysis of mosquito-associated viruses in China aimed to summarize the distribution and diversity of these viruses in the country's provinces and recorded the wide distribution of the virus families among seven mosquito genera (Atoni et al., 2020).

The distribution of similar viruses among different species highlights the potential importance of all mosquito species in an ecosystem to the circulation of mosquito-associated viruses. In the present report, we aim to determine the core virome of all mosquito species present in an ecosystem and analyze the

specificity and distribution of the identified viruses. Comparative analysis of the most abundant viruses present in the ecosystem allowed us to identify distinct signatures of viruses in the different mosquito species and determine those species that may act as virus transmission bridges between genera.

Methods

Mosquito collection and identification

Adult mosquitoes were collected using CDC light traps with CO_2. Collection points spanned across the areas of Eastern Macedonia and Thrace in Greece during two different periods of mosquito activity; April–October 2018 and April– October 2019 (**Supplementary Fig. 1**).

Mosquito specimens were examined over a bed of crushed ice to maintain their condition at all times, both during sample sorting and in making species identifications. Samples were otherwise stored at -80°C prior to RNA and DNA extraction. Female mosquitoes were identified using external morphological features. Species nomenclature follows (Harbach, 2018) and generic abbreviations follow (Wilkerson et al., 2015, appendix S3). Morphological identification was done using a combination of the keys of Saminidou-Voyadjoglou (2001) and Samanidou-Voyadjoglou & Harbach (2011) and the online resource MosKeyTool (Gunay et al., 2018). For the genus *Anopheles* a key to females (Glick, 1992) provided additional morphological characters relevant to the Greek fauna.

Species identification through COI barcoding

Where appropriate, DNA barcoding was done using standard COI PCR and Sanger Sequencing. Mosquitoes were homogenized and total RNA was extracted by TRIzol reagent (Thermo Fischer Scientific) according to the manufacturer's protocol. 1 μg of the RNA extract was reverse-transcribed at 42°C for 60 min, using M-MLV Reverse Transcriptase (Promega). Universal primers COI_F (5' GGATTTGGAAATTGATTAGTTCCTT 3') and COI_R (5' AAAAATTTTAATTCCAGTTGGAACAGC 3') were used to amplify a 600 bp PCR product. The PCR reaction mixture contained 0.25x GC buffer, 1.5 mM $MgCl_2$, 1 mM dNTPs mix, 0.2 μM of each primer and 1.5 U KAPA Taq DNA polymerase (Kapa Biosystems). The thermal profile of the PCR included 40 cycles of denaturation at 95°C for 30 s, annealing at 50°C for 45 s and elongation at 65°C for 1 min, and a final elongation step at 65°C for 7 min. PCR products were purified using the NucleoSpin Gel and PCR Clean-up purification kit (Macherey-Nagel). Sanger Sequencing was performed on the PCR product and analysed using the Barcode of Life Data System V4 platform (Ratnasingham and Hebert, 2007).

Total RNA Next Generation Sequencing

Mosquitoes were homogenized and total RNA was extracted by TRIzol reagent (Thermo Fischer Scientific) according to the manufacturer's protocol. Whole transcriptome libraries were prepared from 500 ng of RNA extract, using the Ion Total RNA-Seq v2 Core Kit (#4479789, ThermoFisher Scientific) according to the manufacturer's instruction. In brief, the RNA library preparation involved RNA

fragmentation, adapter ligation, reverse transcription and 14 cycles of PCR amplification using Ion Xpress™ RNA-Seq Barcode 1–16 Kit (#4475485, ThermoFisher Scientific). Quantification of the library was performed using Qubit Fluorometer high-sensitivity kit (ThermoFisher Scientific) and its median size was determined in LabChip GX Touch 24 (PerkinElmer). The libraries were loaded into an Ion 540 chip, using the automated Ion Chef System (Thermo Fisher Scientific) and sequencing was carried out on an Ion GeneStudio S5, ion torrent sequencer (ThermoFisher Scientific).

Viral genome assembly

Following NGS procedure, raw sequences from the sampled homogenous (each comprised of 5 individuals of the same mosquito species) and heterogenous (mixed mosquito species) pools, were used as input for de novo assembly using Trinity (v2.8.5.) (Grabherr et al., 2011). Trinity assembler, based on the de Bruijn graph algorithm, produces contigs (set of overlapping DNA segments) which represent alternate transcripts of genes, treating sequences with structural changes (mutations and indels) as isoforms of the same gene. The whole process is performed via 3 distinct modules, namely Inchworm, Chrysalis and Butterfly, each respectively responsible for creating the assemblies of transcripts, clustering them and optimizing the de Bruijn graphs. Due to non-deterministic/probabilistic nature of the algorithm, each sample/pool was submitted to 5 Trinity assembly runs, using the default program parameters, thus maximizing the possibilities of revealing *bona fide* full-length viral sequences. The generated assembled contigs of all Trinity runs were aligned against the non-redundant(nr) protein database via BLASTx using taxonomic search restriction on *viridae* (taxid:10239) and annotated by their top BLASTx hit (Altschul et al., 1990). Contigs annotated similarly were fed to the CAP3 online tool for the creation of scaffolds by overlapping contigs, aiming to maximize viral genome assembly efficiency (Huang and Madan, 1999).

Phylogenetic analysis of assembled viruses

For phylogenetic analysis, the protein RNA-dependent RNA polymerase or RdRp (also characterized as segment L or L protein) was used for the construction of the phylogenetic trees. In case RdRp segment was not assembled, nucleocapsid (segment S or N) or glycoprotein (segment M or G) was used. The respective virus amino acid sequence was used as input for BLASTp, and hits with not less than 25% coverage and identity were used for the multiple alignment. Host and geographic origin data were extracted through NCBI tool *Entrez-direct* by employing a custom Python script. Selected sequences were fed to the *NGPhylogeny.fr* website for the elucidation of phylogenetic relationships, running a custom workflow with successive stages of multiple sequence alignment (MAFFT), alignment refinement (BMGE), phylogenetic reconstruction (FastME, based on balanced minimum evolution) for 1000 bootstrap cycles and, finally, graphical representation of the inferred trees (Newick Display) (Lemoine et al., 2019). Trees were exported to iTol (Letunic and Bork, 2019).

Data visualization and statistical analysis

Viruses that presented high amino acid sequence similarity (> 95%) were plotted as links between mosquito species in the habitat and visualized as a Chord Diagram using the circlize package in R (Gu et

al., 2014). Mosquito-related viruses that emerged through BLASTp and presented in the phylogenetic analysis for the 14 families (*Solemoviridae, Luteoviridae, Bunyaviridae, Phasmaviridae, Partitiviridae, Picornaviridae, Chrysoviridae, Totiviridae, Orthomyxoviridae, Rhabdoviridae, Virgaviridae, Narnaviridae, Tymoviridae* and one unclassified family) were plotted against host/carrier mosquito genera using a Sankey Diagram in R implementing the NetworkD3 package (https://www.rdocumentation.org/packages/networkD3/versions/0.4).

Using the quantity of individual species that were included in the pools (**Supplementary Table 1**) and the total number of NGS reads of the respective viruses (species virome), we employed a linear regression model to ascertain if there is predictive value in the carrier/host-virome relationship and determine the stability of core-virome for any given species. The ratio of host/carrier mosquito individuals were plotted against the ratio of the corresponding virus NGS reads between pools (pool A vs. B, C, D, and E) and a linear regression was used in order to assess the hypothesis that individual mosquito species viromes may be predictive of the total mosquito virome.

Results

In order to analyse the species-specific core virome in the ecosystem, mosquitoes were collected during April-October 2019 using CDC traps in rural areas of the prefecture of Eastern Macedonia and Thrace, Greece (14,157 km^2), with a wide variety of landscape features that include sea-shore, mountains (~ 2200 m), lakes, lagoons, swamps, rivers and a large river delta. The prefecture shares borders with Bulgaria and Turkey and in the past has been significantly affected by vector-borne diseases (Groen et al., 2017; Kampen et al., 2002; Nomikou et al., 2009; Papa et al., 2010). Morphological characterisation assisted by cytochrome oxidase subunit I (COI) barcoding resulted in the identification of 24 mosquito species belonging to six genera. Single species pools of five individuals from different sites of the ecosystem were used for total RNA extraction and total RNA next generation sequencing (NGS). Pools of five have been shown in the past to yield all the required data for core virome analysis (Shi et al., 2019). The same procedure was followed for five mixed-species pools of 100 mosquito individuals each that were collected from different places in the same ecosystem (**Supplementary Table 1**).

Through a bioinformatics pipeline that generated sequence contigs and annotated them using virus-restricted BLASTx we were able to identify 35 viruses, including 9 novel virus species with limited phylogenetic relationship to previously known species. In total all 24 mosquito species showcased a wide range of viruses which corresponded to 14 viral families: *Solemoviridae, Luteoviridae, Bunyaviridae, Phasmaviridae, Partitiviridae, Picornaviridae, Chrysoviridae, Totiviridae, Orthomyxoviridae, Rhabdoviridae, Virgaviridae, Narnaviridae, Tymoviridae* and one marked as unclassified (Fig. 1, **black dots**)

Aedes

The genus *Aedes* was the most species rich genus examined in this study and included the following 8 species: *Ae. albopictus, Ae. detritus, Ae. pulcritarsis, Ae. caspius, Ae. sticticus, Ae. vexans, Ae. rusticus* and

Ae. geniculatus. The last three species did not yield any virus sequences, while all the others yielded virus contigs belonging to *Solemoviridae, Luteoviridae, Bunyaviridae, Totiviridae, Orthomyxoviridae, Rhabdoviridae* and *Narnaviridae* families (Fig. 1, Fig. 2).

Two different viruses, Nea Chili luteo-like (Fig. 1, Fig. 2) and Makri bunya-like (Fig. 1, Fig. 2), were detected in *Ae. albopictus* mosquitoes. Possessing a bi-partite genome, the first segment of *Nea Chili luteo-like* aligned optimally to *Wenzhou sobemo-like virus 4*. The first segment of *Nea Chili luteo-like* encoded two separate proteins. The protein encoded second in the segment was identified as RNA-dependent RNA polymerase (RdRp) scored 100% on both BLASTp coverage and identity to Wenzhou sobemo-like virus 4 RdRp (YP_009337376). Similarly, the second segment of Nea Chili luteo-like encoded two separate proteins, with the first one noted as capsid. Each of the encoded proteins however were more divergent but closer to the capsid of Guadeloupe mosquito virus (QEM39258) and the hypothetical protein of Culex inatomii luteo-like virus (BBQ04782) with 69.09% and 37.86% amino acid (aa) identity respectively. The nucleocapsid or segment S of Makri bunya-like virus was detected and successfully assembled, while segments L (RdRp) and M (Glycoprotein) could not be detected in the *Ae. albopictus* samples. The nucleocapsid of Makri Bunya-like virus aligned to Flen bunya-like virus with perfect coverage and only 43.48% identity.

Ae. detritus, carried three different viruses; Evros sobemo-like (Fig. 1, Fig. 2), Ae. detritus orthomyxo-like (Fig. 1, Fig. 2) and Xanthi narna-like (Fig. 1, Fig. 2). Comprised of two segments, which encode two proteins each, Evros Sobemo-like was closest to Atrato sobemo-like virus 1, with the RdRp displaying 90.45% identity to QHA33892, while the capsid scored 77.18% identity against QHA33893. A member of the Orthomyxoviridae family, Aedes detritus orthomyxo-like virus was detected and consisted of four segments, which were predicted to encode three discrete polymerase subunits (PB1, PB2, PA) and a nucleoprotein (NP). All of the above encoded proteins aligned perfectly to the four known segments of Whidbey virus (AQU42764[PB1]/AQU42765[PB2]/AQU42766[PA]/AQU42767[NP]) showing 89.49% [PB1], 92.38% [PB2], 68.42% [PA] and 86.92% [NP] identity. A monopartite virus, named Xanthi narna-like virus, was assembled finally from this species and expressed proteins in both positive and negative strands to resemble Ochlerotatus-associated narna-like virus 2 (AGW51768 and AGW51769). RdRp was predicted to be encoded in the 5'-3' direction, while a hypothetical protein was predicted to be encoded in the 3'-5' direction. Both aligned perfectly to the AGW51768 and AGW51769 database entries (with 98.53% and 90.30% identity respectively).

A similar virus, Xanthi narna-like virus, was found in *Ae. pulcritarsis* but was not assembled completely due to low virus abundance. *Ae. pulcritarsis* mosquitoes contained another member of the sobemo-like viruses, Pythion sobemo-like virus, which aligned best to Atrato sobemo-like virus 4 (QHA33876 and QHA33877), showing 81.84% RdRp identity to QHA33876, while the capsid scored 71.83% identity against QHA33877.

Ae. caspius yielded contigs corresponding to the glycoprotein and the RdRp transcripts of a novel rhabdovirus, Evros Rhabdovirus 1, while the rest of the proteins (Nucleprotein, Matrix, Phosphoprotein)

were fragmented or missing and thus not presented in the overall table (Fig. 1, Fig. 2). Evros Rhabdovirus 1 glycoprotein exhibited 90% coverage but only 27.98% identity with Ictalurid rhabdovirus predicted glycoprotein (AVZ61180). RdRp displayed higher similarity, scoring 100% coverage and 41.33% identity, with Ohlsdorf virus RdRp (ATG83565).

A member of *Totiviridae* family, Drama totivirus (Fig. 1, Fig. 2), was found in *Ae. sticticus* mosquitoes that was related to the Embera virus. Drama totivirus predicted capsid (C) and RdRp shared large coverage with corresponding Embera virus proteins (QHA33711[C]/QHA33712[RdRp]) but with low level of identity (49.33%[C]/51.96%[RdRp]).

Culex

Mosquitoes of the genus *Culex* comprised 5 species; *Cx. modestus, Cx. pipiens, Cx. theileri, Cx. impudicus* and *Cx. perexiguus*. The last three species did not yield any virus sequences, while the others yielded virus contigs belonging to the families *Solemoviridae, Luteoviridae, Virgaviridae, Totiviridae, Partitiviridae, Orthomyxoviridae, Rhabdoviridae* and *Chrysoviridae* (Fig. 1, Fig. 2).

Cx. modestus yielded contigs belonging to 5 viruses, but only 3 viral genomes could be assembled fully ; namely Culex modestus totivirus virus (Fig. 1, Fig. 2), Culex modestus orthomyxo-like virus (Fig. 1, Fig. 2) and Alexandroupolis virga-like virus (Fig. 1, Fig. 2). Dragana partiti-like virus (Fig. 1, Fig. 2) and Soufli chryso-like virus (Fig. 1, Fig. 2) were also detected in fragments. The first segment (RdRp) of Soufli chryso-like virus was aligned successfully against Hubei chryso-like virus 1 (ASA47348) with 89.14% aa identity. Culex modestus totivirus virus was similar to the Culex inatomii totivirus genome, having almost perfect identity (98.20%/98.05%) to the corresponding capsid (C) and RdRp viral database entries (BBQ05101[C]/BBQ05102[RdRp]). Culex modestus orthomyxo-like virus possessed 4 protein segments, predicted to correspond to nucleoprotein (NP) and 3 polymerase subunits (PA, PB1, PB2). All segments were related closely to the Culex orthomyxo-like virus (QGA87316[NP]/QGA87317[PA]/QGA87318[PB1]/QGA87319[PB2]) with high aa identity (86.55% [NP]/93.03%[PA]/96.23%[PB1]/92.22%[PB2]). Alexandroupolis virga-like virus displayed high similarity to Hubei virga-like virus 2 RdRp (YP_009337412) and coat protein (YP_009337413), with 84.04% and 88.78% aa identity respectively.

Cx. pipiens carried a member of the family *Orthomyxoviridae*, Culex pipiens orthomyxo-like virus (Fig. 1, Fig. 2), which was comprised of 6 protein segments encoding 3 polymerase subunits (PB2, PB1, PA), a nucleoprotein (NP), a glycoprotein(G) and a hypothetical protein. The segments corresponded to the *Wuhan Mosquito Virus 6* (ASA47293[PB2]/AJG39094[PB1]/ASA47295[PA]/ASA47477[NP]/ASA47297[G]/ASA47298[hypothetical protein]), with 100% coverage and above 97 % aa identity.

In the Cx. theileri pool, the novel Thassos sobemo-like virus (Fig. 1, Fig. 2) was detected along with a virga-like virus in fragments. Thassos sobemo-like virus was assembled completely, showing limited similarity in predicted RdRp and capsid protein to Atrato sobemo-like virus 1 (QHA33892) and Bat

sobemovirus (AGN73380) respectively. The RdRp of the novel Thassos Sobemo-like virus resembled the Atrato Sobemo-like virus 1 RdRp with 53.44% aa identity, while the capsid (C) and hypothetical protein products of Thassos Sobemo-like virus best matched bat sobemovirus (AGN73380[C]/BBQ04498[hyp.protein]), showing limited aa identity (36.51%[C]/26.09%[hyp.protein]).

A virus predicted to belong to the family *Luteoviridae*, Culex impudicus luteo-like virus (Fig. 1, Fig. 2), was detected in *Cx. impudicus* and resembled Culex luteo-like virus. Each of the encoded proteins, including a highly conserved RdRp and capsid(C), aligned perfectly to Culex luteo-like database entries ASA47391[hyp.protein]/ ASA47392[RdRp]/ ASA47393[C]/ ASA47394[hyp.protein] with 100% coverage and above 82% of aa identity. Despite the detection of a predicted nucleocapsid (segment S) of Culex phasma-like virus (Fig. 1, Fig. 2) in *Cx. impudicus* mosquitoes, it was assembled partially and both segments L (RdRp) and M (Glycoprotein) were missing. Culex phasma-like virus nucleocapsid had 85.50% aa identity with Terena virus segment S (ANW72255).

Cx. perexiguus, carried Culex perexiguus rhabdovirus (Fig. 1, Fig. 2) from which only the glycoprotein portion could be assembled successfully, showing 79.09% identity to Culex mononega*-like virus 1* glycoprotein (ASA47286).

Culiseta

Two species belonging to the genus *Culiseta* were identified; *Cs. longiareolata* and *Cs. annulata*. *Cs. longiareolata* carried 2 viruses belonging to different families, Soufli partiti-like virus (Fig. 1, Fig. 2) and Evros chryso-like virus (Fig. 1, Fig. 2). Soufli partiti-like virus was assembled fully resembling the Atrato partiti-like virus 2 RdRp (QHA33902) and capsid (QHA33903) segments with significant aa identity at 82.16% and 51.07% respectively. Evros chryso-like virus RdRp segment was assembled completely and presented high coverage and identity to Shuangao chryso-like virus 1 RdRp (QDC23188, 98% coverage and 70% identity). A novel bunyavirus, Evros bunya-like virus, was detected within *Cs annulata* mosquitoes (Fig. 1, Fig. 2), which deviated from known homologous viral sequences. The nucleocapsid (segment S) of Evros bunya-like virus aligned to a *Bunyaviridae* family contig from an environmental sample metagenome (QIN93566), displaying 98% coverage and 38.34% identity, while the glycoprotein (segment M) was closest to the Culex pseudovishnui bunya-like virus database entry BBQ05094, with 96% coverage and 34.99% identity. Notably, RdRp (segment L) of the novel Evros bunya-like virus could not be assembled, possibly due to low abundance.

Coquillettidia

Two mosquito species belonging to the genus *Coquillettidia* were sampled, *Cq. buxtoni* and *Cq. richardii. Cq. buxtoni* yielded no virus contigs, while *Cq. richardii* harbored viruses of the families *Bunyaviridae* and *Partitiviridae* and an unclassified virus. The nucleocapsid (segment S) of the novel Abdera Bunya-like virus (Fig. 1, Fig. 2) was detected and assembled. The assembled genome lacked RdRp and glycoprotein segments, possibly due to low abundance. The nucleocapsid of Abdera Bunya-like virus exhibited 95% coverage and 36.09% identity against Culex pseudovishnui bunya-like virus segment S (BBU41515). Xanthi partiti-like virus (Fig. 1, Fig. 2) was assembled successfully in *Cq. Richardii*, possessing separate

RdRp and capsid protein segments. These proteins resembled the corresponding partiti-like virus 1 RdRp (AOR51388, 89% coverage and 77.95% identity) and the Atrato partiti-like virus 2 capsid (QHA33903, 99% coverage and 51.66% identity). A previously unclassified virus, distinguished from other characterized virus families, was found within *Cq. richardii* mosquitoes, here named Didymoteicho Chaq virus (Fig. 1, Fig. 2), which was related closely to Chaq virus-like 1 (AOR51384) with perfect coverage and 82.77% identity.

Anopheles

The genus *Anopheles* was the second most species rich group of mosquitoes examined and consisted of 6 species, *An. melanoon, An. pseudopictus, An. sacharovi, An. claviger, An. algeriensis and An. superpictus.* Sequencing of these species yielded virus contigs belonging to the families *Solemoviridae, Luteoviridae, Partitiviridae and Rhabdoviridae* (Fig. 1, Fig. 2).

An. melanoon carried a partiti-like virus, Dragana partiti-like virus (Fig. 1, Fig. 2), whose segments aligned best to the partitivirus-like 2 RdRp (AOR51389) and Beihai partiti-like virus 2 putative capsid (YP_009333351). Specifically, the RdRp segment showed 96% coverage and 70.71% aa identity to the viral database entry AOR51389, while the capsid had 81% coverage and 33.83% aa identity to the YP_009333351 database entry.

Two closely related members of the family *Bunyaviridae* were detected separately in the species *An. pseudopictus* and *An. superpictus*, and these were related to Culex bunya-like virus RdRp (AXQ04766) and nucleocapsid (AXQ04767) segments respectively. Both viruses were 100% covered by the corresponding viral database entries, with the first one, Rhodopi bunya-like virus (Fig. 1, Fig. 2) showing 71.01% aa identity, while Anopheles bunya-like virus exhibited much greater identity (98.80%) (Fig. 1, Fig. 2).

The novel Xanthi rhabdovirus (Fig. 1, Fig. 2) was found in *An. sacharovi* mosquitoes and was assembled successfully as a complete genome, possessing all of its 5 sequentially transcribed proteins, annotated as N, M, P, G and L proteins. G protein of Xanthi rhabdovirus was 26.42% identical to the Marco virus (YP_009362114), while and L protein was 56.03% identical to the Beaumont virus (AGX86091).

An. algeriensis carried 3 novel viruses. The novel Ferres luteo-like virus (Fig. 1, Fig. 2) resembling Sanxia sobemo-like virus 3. The encoded proteins of Ferres Luteo-like virus displayed over 93% coverage and 39.96%[hyp.protein 1]/66.04%[hyp.protein 2] aa identity to the corresponding database entries of Sanxia sobemo-like virus 3 (YP_009336905[hyp.protein 1]/YP_009336906[hyp.protein 2]). Another novel virus within the *An. algeriensis* pool was Evros rhabdovirus 2 (Fig. 1, Fig. 2) whose genome was almost assembled fully and, apart from its M protein coding region. Evros rhabdovirus 2, showed high rates of coverage to the respective viral database entries, aligning best to sequences of various members of the family *Rhabdoviridae* such as the Obodhiang virus nucleoprotein or N protein (YP_006200957, 28.04% identity), the Culex rhabdo-like virus glycoprotein or G protein (AXQ04773, 23.11% identity) and the Merida virus RdRp or L protein (AWJ96718, 43.11% identity). Finally, a novel partiti-virus, Orestiada partiti-

like virus was assembled partially (Fig. 1, Fig. 2). Only the capsid protein segment of Orestiada partiti-like virus was assembled successfully and aligned best on the Beihai partiti-like virus 2 capsid (YP_009333351) with 56% coverage and 35.54% aa identity. A similar virus was also detected in the *An. claviger* pool, although, it was significantly fragmented.

Uranotaenia

Despite being the smallest genus examined, both in species number and individual physical size, *Uranotaenia* and its only representative, *Ur. unguiculata*, harbored five viruses belonging to the families *Phasmaviridae*, *Virgaviridae*, *Chrysoviridae*, *Totiviridae* and *Tymoviridae*. Two of these viruses, Alexandroupolis virga-like virus (Fig. 1, Fig. 2) and Uranotaenia phasma-like virus (Fig. 1, Fig. 2), had mainly fragmented genomic sequences and therefore could not be assembled fully. However, nucleocapsid (segment S) of Uranotaenia phasma-like virus was assembled successfully, displaying 100% coverage and 85.10% identity to the Culex phasma-like virus nucleocapsid (ASA47280). Another virus similar to Soufli chryso-like virus, named Xanthi chryso-like virus (Fig. 1, Fig. 2), was found within *Ur. unguiculata*. Each of the 4 protein segments constituting Xanthi chryso-like virus were assembled completely, showing almost complete similarity to Hubei chryso-like virus 1 (ASA47348) with identity ranging between 75% and 90% depending on the segment. A novel member of the family *Totiviridae*, the Xanthi Totivirus (Fig. 1, Fig. 2), was also detected. The encoded capsid and RdRp proteins of Xanthi Totivirus resembled the corresponding proteins of the Australian Anopheles totivirus (QIJ25845) and Pisingos virus (QHA33716), showing perfect coverage with 40.65% and 45.01% identity respectively. Lastly, a novel tymovirus, Xanthi tymo-like virus (Fig. 1, Fig. 2), was assembled. Xanthi tymo-like virus RdRp was related to the Culex pseudovishnui tymo-like virus RdRp (BBQ04465) with 88% coverage and 54.96% identity.

Mosquito pools

The same procedure, as described above for the individual species, was performed on the five pools of mixed mosquito species. The analysis returned the vast majority of BLASTx hits that were present in the individual species, except two (Fig. 1, **black dots**). In pools C and D two unrelated viruses belonging to the family *Picornaviridae* were detected that had not been previously identified in the samples of individual species. These viruses, Thrace picorna-like 1 virus and the Thrace picorna-like 2 virus, were assembled at full length due to the high abundance of the respective NGS reads. The encoded polyproteins of Thrace picorna-like 1 virus and Thrace picorna-like 2 virus were related closely to Yongsan picorna-like virus 1 (AXV43880, 74.32% aa identity) and Flen picorna-like virus (QGA87323, 71.59% aa identity) respectively.

Host-virus networks in the ecosystem

The viruses that were identified and assembled from the individual species, together with the five pools of mixed species, were used to build a reference database of core mosquito viromes in the ecosystem. NGS reads from all individual species and the five pools of mixed species were aligned against the reference database for the identification viruses of low abundance or viruses missed during the assembly-BLASTx

pipeline. The revised table of viruses in individual species (Fig. 1, **heatmap**) had 92% overlap with the initial assembly-BLASTx table (Fig. 1, **black dots**). The situation was completely different in the mixed-species populations (pools) as the use of the ecosystem-specific virus reference revealed the presence of low abundance viral signatures across the pools.

In order to build an ecosystem-based host-virus network we compared the core viromes of all species to identify common paths of virus transmission. Analysis the core viromes of individual species showed that certain viruses were shared, not only between species of the same genus, but also between different genera (Fig. 3). As expected most of the common viruses were shared between species of the same genus. *Ae. detritus* and *Ae. pulcritarsis* shared two variants of Xanthi narna-like virus with 89.74% nucleotide similarity (Table 1). *An. algeriensis* and *An. claviger* shared two variants of Orestiada partiti-like virus segment 2 with 91.5% nucleotide identity (Table 1). Rhodopi bunya-like virus and Anopheles bunya-like virus segments found in *An. superpictus* and *An. pseudopictus*, aligned independently to the two segments of a previously identified virus (AXQ04766, AXQ04767) and although they do not share any common segment, were considered similar (Table 1).

Table 1

Nucleotide identity among viruses present in more than one mosquito species. Higher abundance species are those species where full-length viruses were assembled while lower abundance species are those species that yielded partial sequences. Nucleotides (nts) overlap is the number of nucleotides aligned between full-length and partial sequences. The nucleotide percent identity is also reported.

Virus	Higher abundance species	Lower abundance species	Nts overlap	Identity (%)
Xanthi narna-like	*Ae. detritus*	*Ae. pulcritarsis*	692	89.74
Orestiada partiti-like	*An. algeriensis*	*An. claviger*	459	91.50
Dragana partiti-like segment 2	*An. melanoon*	*Cx. modestus*	839	98.57
Alexandroupolis virga-like	*Cx. modestus*	*Ur. unguiculata*	2068	90.33
Alexandroupolis virga-like	*Cx. modestus*	*Cx. theileri*	6646	87.35
Alexandroupolis virga-like	*Ur. unguiculata*	*Cx. theileri*	6647	87.38
Culex phasma-like	*Ur. unguiculata*	*Cx. impudicus*	2107	59.85
Xanthi / Soufli chryso-like segment 1	*Ur. unguiculata*	*Cx. modestus*	3533	87.74
Xanthi / Soufli chryso-like segment 2	*Ur. unguiculata*	*Cx. modestus*	2998	85.69
Xanthi / Soufli chryso-like segment 3	*Ur. unguiculata*	*Cx. modestus*	2893	82.99
Xanthi / Soufli chryso-like segment 4	*Ur. unguiculata*	*Cx. modestus*	3078	81.74

Interestingly, *Cx. modestus* shared common viruses with a variety of species including *Cx. theileri* (Alexandroupolis virga-like virus, 87.35% nucleotide identity), *An. melanoon* (Dragana partiti-like virus segment 2, 98.57% nucleotide identity) and *Ur. unguiculata* (Xanthi / Soufli chryso-like virus, 84,54% nucleotide identity [mean]; Alexandroupolis virga-like virus, 90.33% nucleotide identity) (Table 1). *Ur. unguiculata* emerged also as an important hub of viruses as, apart from the two viruses shared with *Cx. modestus*, it shared also the Alexandroupolis virga-like virus with *Cx. theileri* (87.38%) and the Culex phasma-like virus with *Cx. impudicus*. However, the nucleotide identity with the virus present in *Cx. impudicus* dropped to 59.84%, but with 92.59% aa identity (Table 1).

Novel links between viral families and mosquito genera.

As representatives of the genera *Culiseta* and *Uranotaenia* together with a multitude of species within the genera *Aedes*, *Culex*, *Anopheles* and *Coquillettidia* were sequenced for the first time we attempted to identify novel links between virus families and mosquitoes. Host-annotated phylogenetic trees of the viruses presented in this study provided data on the associations between mosquito genera and the virus families they carry (Fig. 4). Eleven new connections in the global host-virus network at the genus level were identified in this study. This finding highlights the fact that multiple mosquito species may contribute to the circulation of the core virome in any given ecosystem.

Viromes of individual species are representative of the ecosystem mosquito-virome

As core viromes are mostly stable within species, it was hypothesized that the composition of the mosquito virome in the ecosystem would fluctuate in accordance with the composition of the mosquito population. After mapping the NGS reads from the pools of mixed mosquito species to the ecosystem virus reference, we quantified the read count for each virus per pool. Given that the mixed mosquito pools were of known species composition, we were able to correlate virus abundance (read counts) with the abundance of the respective species. The viromes for most species were unique, as noted in the literature and this study. When more than one virus was hosted/carried by a species the sum of virus reads was used. Plotting virus reads against the number of individuals of a species across different pools we observed significant correlation between the abundance of mosquito species and the abundance of the respective viruses that form their core virome (Fig. 5).

To assess the hypothesis that species viromes may be predictive of the total mosquito virome of the ecosystem we plotted the ratio of host/carrier mosquito individuals against the ratio of the corresponding virus/es (individual core virome) reads between pools (pool A vs. B, C, D, and E). Linear regression showed good correlation between reads and individuals (Fig. 5). The analysis highlighted the importance of individual species of an ecosystem in defining the total virus-host/carrier equilibrium.

Discussion

During the past decade, a great number of mosquito species have been analyzed with metagenomics methodologies for the identification of medically important viruses and/within their core virome. Previous works focused on a limited number of species, usually one or two, in order to characterize in depth, or compare their core viromes in a specific ecosystem. In our report, we sought to analyze globally the mosquito core virome of all mosquitoes occurring and identified in an ecosystem and compare their viromes, seeking divergence and convergence points.

Taking into consideration the evolutionary or current cross talk between species viromes and the fact that multiple mosquito species may play complementary roles in the an ecosystem, at an either an ecological or public health level, we analyzed the global mosquito virome of a localized ecosystem. Thrace is situated on the southeastern border of Europe with Asia and between the mountain ranges of the Balkan Mountains and the Aegean Sea. Lakes, lagoons, swamps, rivers and a large river delta characterize this small geographic region that encompasses a high level of biodiversity (Legakis et al., 2018). Indeed, we were able to analyze representatives of 24 different mosquito species, all of which were encountered, in varying ratios, in different parts of the local ecosystem.

In depth analysis of the core viromes of 24 mosquito species revealed the abundant presence of 35 viruses that belong to 14 virus families. Most of the viruses were related to previously described isolates. The viruses identified as sobemo-like viruses, luteo-like viruses, orthomyxoviruses, picornaviruses, chryso-like viruses, virga-like viruses and narna-like displayed more than 85% similarity to known viruses (Fig. 2). On the other hand 9 viruses showed < 55% amino acid similarity to their best match in the NCBI database. Two novel bunyaviruses (Abdera and Evros bunya-like) were identified from the mosquito species *Cq. richardii* and *Cs. annulata*, both of which formed a distinct group within the bunyaviruses tree, although they were located within the branch related to the mosquito genus *Culex* (Faizah et al., 2020; Sadeghi et al., 2018). The identification of a bunyavirus in the mosquito genus *Culiseta*, although never reported as a metagenomics finding, has been reported in the past (McLean et al., 1977). A mosquito species belonging to the genus *Uranotaenia* was analyzed for the first time, yielding both a novel totivirus (Xanthi totivirus) that was located in a cluster of mosquito-related totiviruses previously reported in the mosquito genera *Coquillettidia*, *Anopheles* and *Culex* (Colmant et al., 2017; Faizah et al., 2020; Pettersson et al., 2019) and a novel tymo-like virus located in a distinct cluster of *Culex*-related viruses (Faizah et al., 2020; Pettersson et al., 2019; Shi et al., 2019). Another novel totivirus (Drama totivirus) was identified from the mosquito species *Ae. sticticus* that clustered together with other mosquito-related totiviruses found in the mosquito genera *Aedes*, *Anopheles* and *Wyeomyia* (Fauver et al., 2016; Shi et al., 2019). Interestingly, a novel partitivirus (Orestiada partitivirus), identified in the mosquito species *An. algeriensis*, formed an outgroup in the phylogenetic tree that was located closer to viruses from other insects. Finally, three novel rhabdoviruses (Xanthi rhabdovirus, Evros rhabdovirus 1 and 2) were identified in mosquito species of two genera, *Ae. caspius, An. algeriensis* and *An. sacharovi*. These viruses formed a distinct subgroup that encompassed only one previously known virus; the Beaumont virus from the mosquito *Anopheles annulipes* (Coffey et al., 2014). Although there is lack of data on related rhabdoviruses one may speculate that these viruses, despite the big difference in nucleotide identity, may form a local evolutionary branch.

The core viromes identified here were very distinct among mosquito species, although certain viruses such as Xanthi narna-like, Orestiada partiti-like, Dragana partiti-like, Alexandroupolis virga-like and Xanthi / Soufli chryso-like viruses were very similar between mosquito species of the same genus or even between species of different mosquito genera. Despite the fact that some viruses were shared between different mosquito species, they were substantially divergent (~ 60–90%) in nucleotide sequence. This difference in similarity may argue against virus sharing between mosquito species in the habitat, but may rather indicate a recent adaptation event (Grubaugh et al., 2016). As an exception Dragana partiti-like virus, shared between *An. melanoon* and *Cx. modestus*, showed almost complete similarity between the different mosquito species and genera, despite their evolutionary distance. Core virome genomes are often conserved and stable in either relatively small places, such as Guadeloupe, or large geographical areas such as Australia, Southeast Asia and Northern Europe (Pettersson et al., 2019; Shi et al., 2019; Shi et al., 2017). The stability of the core virome has been shown recently between individual members of *Ae. aegypti* and *Cx. quinquefasciatus* occurring in the same ecosystem, where most of the individuals encompassed the same set of viral genomes (Shi et al., 2019). Small differences in a mosquito species core virome were recorded in different habitats of Northern Europe (Pettersson et al., 2019). The core virome seems to characterize the individual mosquito species, despite overlaps between species of the same genus (Pettersson et al., 2019; Shi et al., 2017), while there is still an open question as to whether these viromes differ significantly between different habitats within a given ecosystem.

Assigning a specific virome to a specific mosquito species and hypothesizing that the core virome may remain stable between different habitats in an ecosystem, we analyzed the virome of population pools mixed species of mosquitoes from different areas of Thrace. A correlation analysis between the number of individuals of a mosquito species and the abundance of the species-specific virome in the populations of five pools of mixed mosquito species showed a linear regression of high statistical significance (p = 0.00017). Furthermore, by plotting the fluctuation of the number of individuals of every species with the fluctuation of each respective species-specific virome, we were able to observe the same effect in the Thrace ecosystem as a whole. This linear regression analysis supported the hypothesis that a stable species-specific mosquito virome exists in the different habitats of an ecosystem and that this virome is defined by the specific composition of the mosquito population.

Conclusions

Our study aimed to assess globally the core mosquito virome in a diverse ecosystem. We discovered 35 viruses, 9 of which were novel, while novel relationships between mosquito species and virus families were identified. These novel relationships mainly rose as a consequence of RNA sequencing for the first time of a wide range of different species that had not been analyzed in the past. Our data further supported the idea of a stable core virome, characteristic of each mosquito species. The remarkable stability of the core virome seemed to determine the composition of the total mosquito core virome of a habitat in the ecosystem, directly related to the local mosquito population.

Abbreviations

aa, amino acid

CDC, Center of Disease Control

COI, cytochrome oxidase subunit I

NGS, next generation sequencing

nts, nucleotides

RdRp, RNA-dependent RNA polymerase

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The raw sequencing datasets for the current study are available in the NCBI Sequence Read Archive (SRA) repository, under the Bioproject with accession code PRJNA681030 (www.ncbi.nlm.nih.gov/bioproject/681030). NCBI GenBank accession numbers for the viruses identified in the study can be found on **Supplementary Table 2**.

Competing interests

The authors declare that they have no competing interests

Funding

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T1EDK-5000).

Author Contributions

IK designed the study. IK and SV obtained funding for the project. AN conducted fieldwork and collected samples. MGRF and MdCW performed morphological species identification. KoK, AK, KaK performed

experiments. KK and ND performed the analysis. KK, ND and IK wrote the manuscript with input from all authors.

Acknowledgements

We would like to thank Professor Penelope Mavromara for helpful discussions.

References

Agboli, E., Leggewie, M., Altinli, M., Schnettler, E., 2019. Mosquito-Specific Viruses-Transmission and Interaction. Viruses 11.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of molecular biology 215, 403-410.

Atoni, E., Zhao, L., Hu, C., Ren, N., Wang, X., Liang, M., Mwaliko, C., Yuan, Z., Xia, H., 2020. A dataset of distribution and diversity of mosquito-associated viruses and their mosquito vectors in China. Scientific data 7, 342.

Coffey, L.L., Page, B.L., Greninger, A.L., Herring, B.L., Russell, R.C., Doggett, S.L., Haniotis, J., Wang, C., Deng, X., Delwart, E.L., 2014. Enhanced arbovirus surveillance with deep sequencing: Identification of novel rhabdoviruses and bunyaviruses in Australian mosquitoes. Virology 448, 146-158.

Colmant, A.M.G., Etebari, K., Webb, C.E., Ritchie, S.A., Jansen, C.C., van den Hurk, A.F., Bielefeldt-Ohmann, H., Hobson-Peters, J., Asgari, S., Hall, R.A., 2017. Discovery of new orbiviruses and totivirus from Anopheles mosquitoes in Eastern Australia. Archives of virology 162, 3529-3534.

Faizah, A.N., Kobayashi, D., Isawa, H., Amoa-Bosompem, M., Murota, K., Higa, Y., Futami, K., Shimada, S., Kim, K.S., Itokawa, K., Watanabe, M., Tsuda, Y., Minakawa, N., Miura, K., Hirayama, K., Sawabe, K., 2020. Deciphering the Virome of Culex vishnui Subgroup Mosquitoes, the Major Vectors of Japanese Encephalitis, in Japan. Viruses 12.

Fauver, J.R., Grubaugh, N.D., Krajacich, B.J., Weger-Lucarelli, J., Lakin, S.M., Fakoli, L.S., 3rd, Bolay, F.K., Diclaro, J.W., 2nd, Dabire, K.R., Foy, B.D., Brackney, D.E., Ebel, G.D., Stenglein, M.D., 2016. West African Anopheles gambiae mosquitoes harbor a taxonomically diverse virome including new insect-specific flaviviruses, mononegaviruses, and totiviruses. Virology 498, 288-299.

Goenaga, S., Kenney, J.L., Duggal, N.K., Delorey, M., Ebel, G.D., Zhang, B., Levis, S.C., Enria, D.A., Brault, A.C., 2015. Potential for Co-Infection of a Mosquito-Specific Flavivirus, Nhumirim Virus, to Block West Nile Virus Transmission in Mosquitoes. Viruses 7, 5801-5812.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren,

B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29, 644-652.

Groen, T.A., L'Ambert, G., Bellini, R., Chaskopoulou, A., Petric, D., Zgomba, M., Marrama, L., Bicout, D.J., 2017. Ecology of West Nile virus across four European countries: empirical modelling of the Culex pipiens abundance dynamics as a function of weather. Parasites & vectors 10, 524.

Grubaugh, N.D., Weger-Lucarelli, J., Murrieta, R.A., Fauver, J.R., Garcia-Luna, S.M., Prasad, A.N., Black, W.C.t., Ebel, G.D., 2016. Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching. Cell host & microbe 19, 481-492.

Gu, Z., Gu, L., Eils, R., Schlesner, M., Brors, B., 2014. circlize Implements and enhances circular visualization in R. Bioinformatics 30, 2811-2812.

Hall-Mendelin, S., McLean, B.J., Bielefeldt-Ohmann, H., Hobson-Peters, J., Hall, R.A., van den Hurk, A.F., 2016. The insect-specific Palm Creek virus modulates West Nile virus infection in and transmission by Australian mosquitoes. Parasites & vectors 9, 414.

Hobson-Peters, J., Yam, A.W., Lu, J.W., Setoh, Y.X., May, F.J., Kurucz, N., Walsh, S., Prow, N.A., Davis, S.S., Weir, R., Melville, L., Hunt, N., Webb, R.I., Blitvich, B.J., Whelan, P., Hall, R.A., 2013. A new insect-specific flavivirus from northern Australia suppresses replication of West Nile virus and Murray Valley encephalitis virus in co-infected mosquito cells. PloS one 8, e56534.

Huang, X., Madan, A., 1999. CAP3: A DNA sequence assembly program. Genome research 9, 868-877.

Kampen, H., Maltezos, E., Pagonaki, M., Hunfeld, K.P., Maier, W.A., Seitz, H.M., 2002. Individual cases of autochthonous malaria in Evros Province, northern Greece: serological aspects. Parasitology research 88, 261-266.

Kenney, J.L., Solberg, O.D., Langevin, S.A., Brault, A.C., 2014. Characterization of a novel insect-specific flavivirus from Brazil: potential for inhibition of infection of arthropod cells with medically important flaviviruses. The Journal of general virology 95, 2796-2808.

Legakis, A., Constantinidis, T., Petrakis, P.V., 2018. Biodiversity in Greece. Global biodiversity. Volume 2, Selected countries in Europe, 71-124.

Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., Gascuel, O., 2019. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. Nucleic acids research 47, W260-W265.

Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic acids research 47, W256-W259.

Li, C.X., Shi, M., Tian, J.H., Lin, X.D., Kang, Y.J., Chen, L.J., Qin, X.C., Xu, J., Holmes, E.C., Zhang, Y.Z., 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. eLife 4.

McLean, D.M., Grass, P.N., Judd, B.D., Ligate, L.V., Peter, K.K., 1977. Bunyavirus isolations from mosquitoes in the western Canadian Arctic. The Journal of hygiene 79, 61-71.

Nasar, F., Erasmus, J.H., Haddow, A.D., Tesh, R.B., Weaver, S.C., 2015. Eilat virus induces both homologous and heterologous interference. Virology 484, 51-58.

Nomikou, K., Dovas, C.I., Maan, S., Anthony, S.J., Samuel, A.R., Papanastassopoulou, M., Maan, N.S., Mangana, O., Mertens, P.P., 2009. Evolution and phylogenetic analysis of full-length VP3 genes of Eastern Mediterranean bluetongue virus isolates. PloS one 4, e6437.

Papa, A., Dalla, V., Papadimitriou, E., Kartalis, G.N., Antoniadis, A., 2010. Emergence of Crimean-Congo haemorrhagic fever in Greece. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 16, 843-847.

Pettersson, J.H., Shi, M., Eden, J.S., Holmes, E.C., Hesson, J.C., 2019. Meta-Transcriptomic Comparison of the RNA Viromes of the Mosquito Vectors Culex pipiens and Culex torrentium in Northern Europe. Viruses 11.

Ratnasingham, S., Hebert, P.D., 2007. bold: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular ecology notes 7, 355-364.

Sadeghi, M., Altan, E., Deng, X., Barker, C.M., Fang, Y., Coffey, L.L., Delwart, E., 2018. Virome of >12 thousand Culex mosquitoes from throughout California. Virology 523, 74-88.

Schultz, M.J., Frydman, H.M., Connor, J.H., 2018. Dual Insect specific virus infection limits Arbovirus replication in Aedes mosquito cells. Virology 518, 406-413.

Shi, C., Beller, L., Deboutte, W., Yinda, K.C., Delang, L., Vega-Rua, A., Failloux, A.B., Matthijnssens, J., 2019. Stable distinct core eukaryotic viromes in different mosquito species from Guadeloupe, using single mosquito viral metagenomics. Microbiome 7, 121.

Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., Qin, X.C., Li, J., Cao, J.P., Eden, J.S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., Zhang, Y.Z., 2016. Redefining the invertebrate RNA virosphere. Nature 540, 539-543.

Shi, M., Neville, P., Nicholson, J., Eden, J.S., Imrie, A., Holmes, E.C., 2017. High-Resolution Metatranscriptomics Reveals the Ecological Dynamics of Mosquito-Associated RNA Viruses in Western Australia. Journal of virology 91.

Vasilakis, N., Tesh, R.B., 2015. Insect-specific viruses and their potential impact on arbovirus transmission. Current opinion in virology 15, 69-74.

Figures

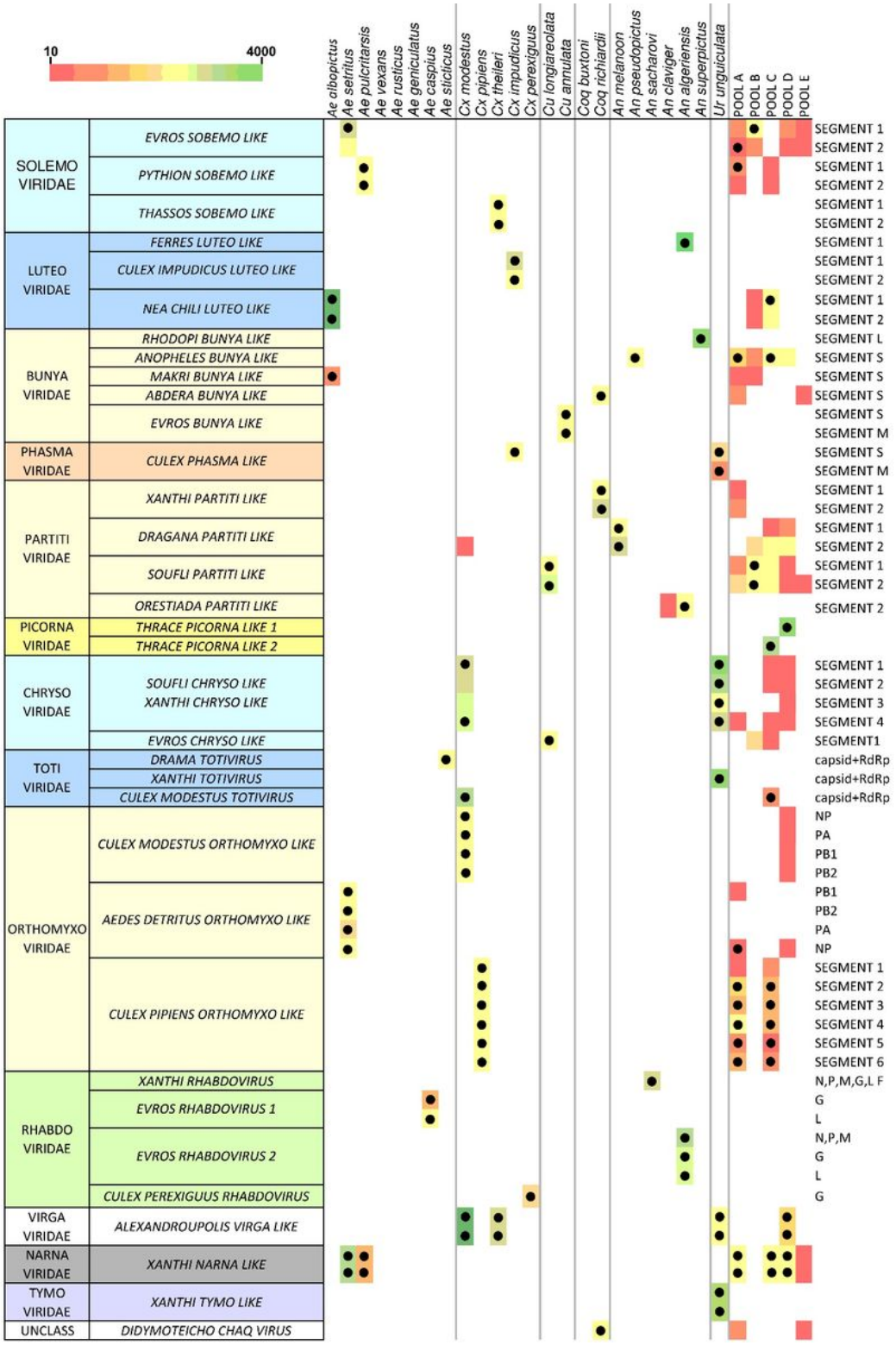

Figure 1

Table of viruses identified in the present report. The viruses are grouped according to the family in which they were predicted to be classified. Viruses in the 24 individual mosquito species and the 5 pools are reported. For segmented viruses, segment names are noted on the right. Black dots represent the identification of virus contigs after assembly-BLASTx. Heatmap is a heatmap of the quantity of NGS reads after alignment against the regional virus database. The color gradient represents the range of the counted NGS reads per virus segment.

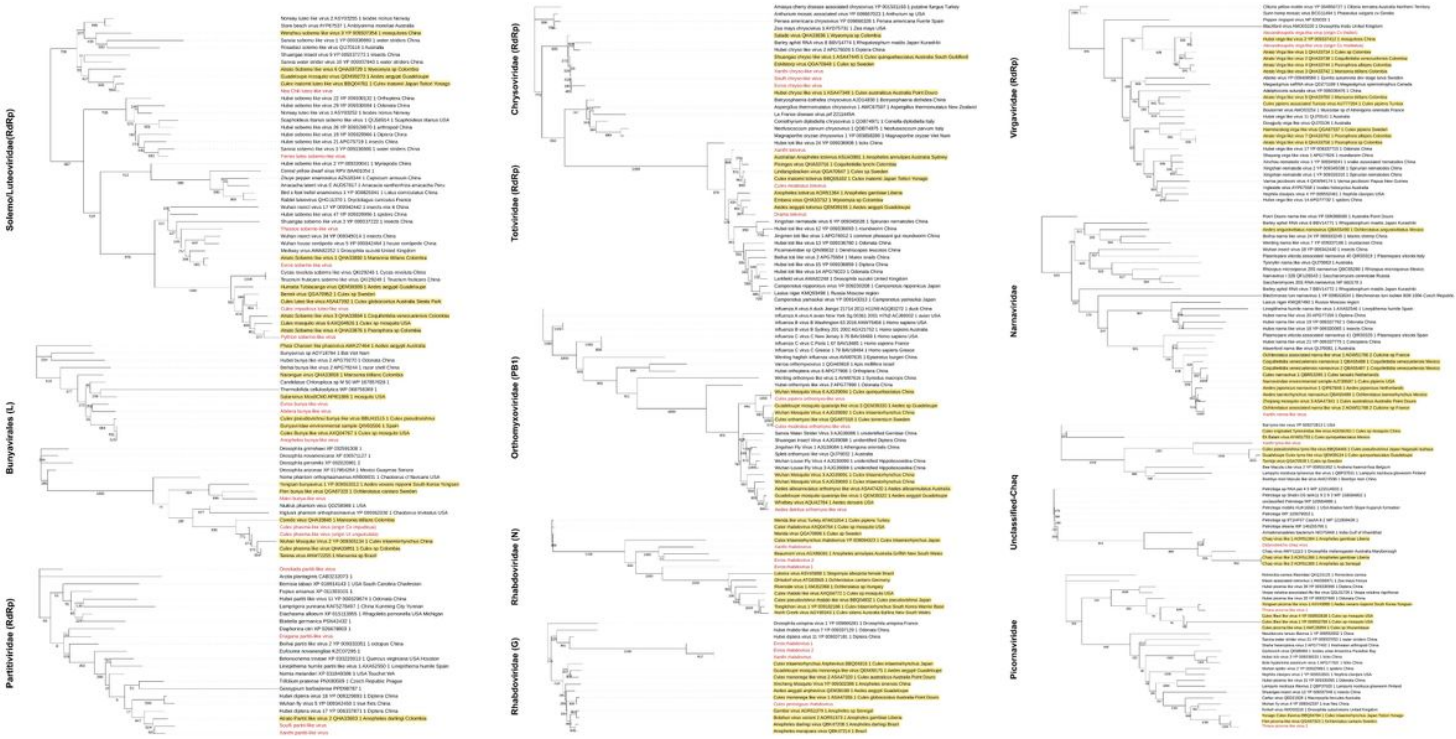

Figure 2

Phylogenetic trees for the phylogenetic analysis of the identified viruses in the study. RNA-dependent RNA polymerase (RdRp) was used for the construction of the phylogenetic trees. In cases where the RdRp segment was not assembled, nucleocapsid (N) or glycoprotein (G) was used. Trees were constructed in NGPhylogeny.fr using MAFFT alignment, BMGE alignment refinement and FastME. Bootstrap values, indicated at the nodes, were obtained from 1000 bootstrap replicates. Viruses highlighted in yellow have been identified in mosquito species. Viruses in red font are the viruses identified in the present study

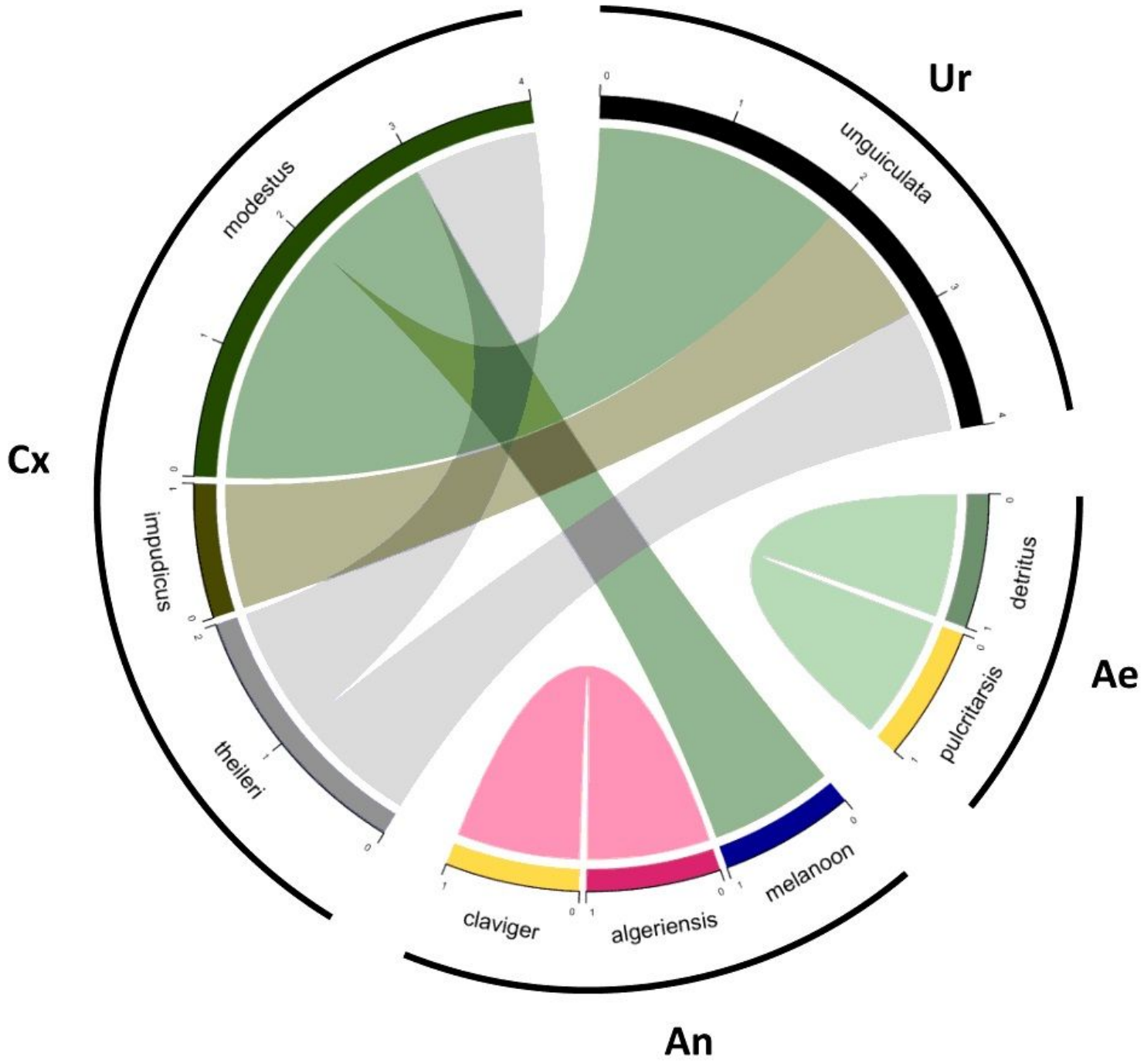

Figure 3

Virus-carrier/host Chord Diagram network in the Thrace ecosystem. Viruses that presented high amino acid sequence similarity (>95%) were plotted as links between mosquito species in the habitat.

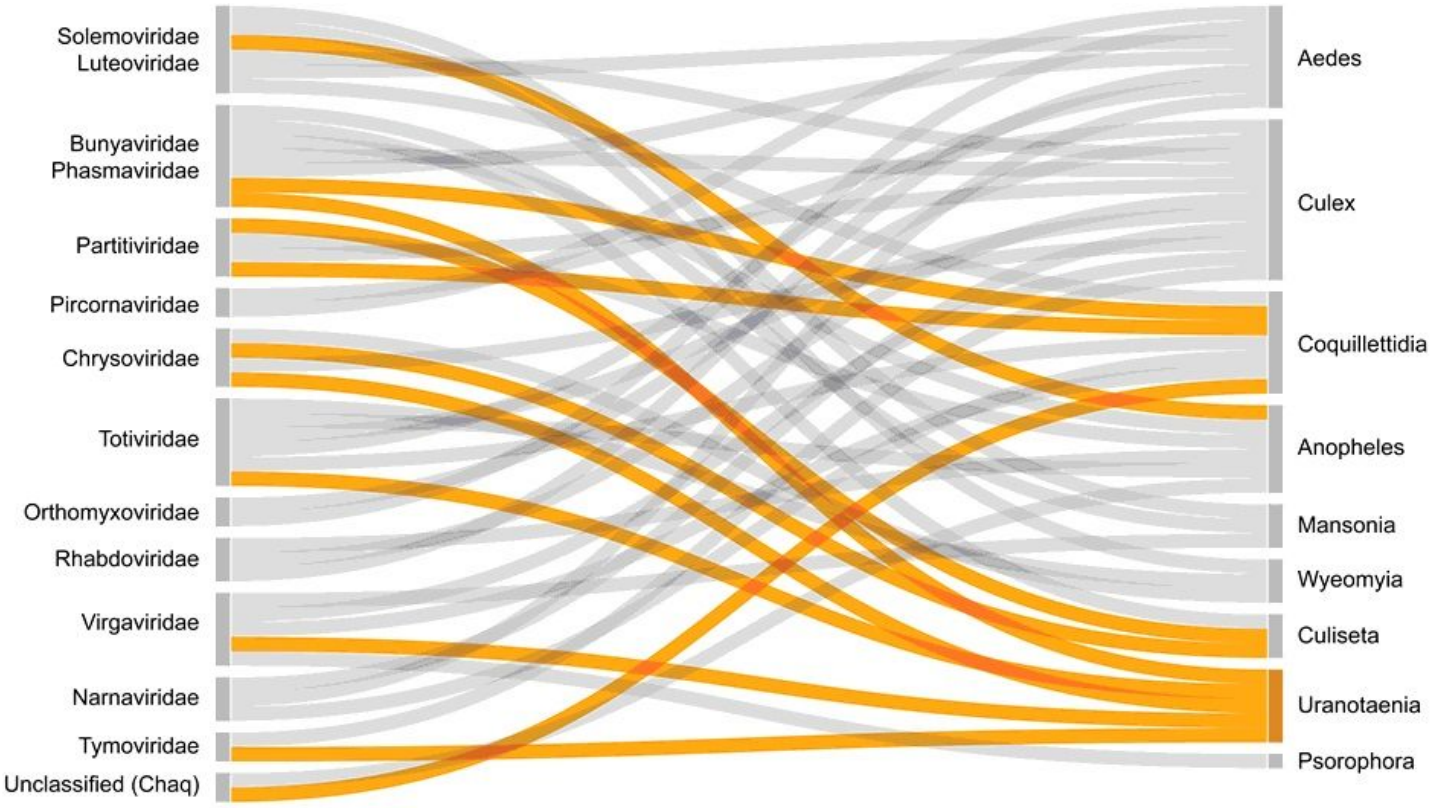

Figure 4

Sankey diagram of the global virus-carrier/host network involving viruses with >25% amino acid similarity to the viruses identified in the study. On the left side are the families of viruses and on the right side are the respective carrier/host mosquito genera. Grey chords represent links reported in the literature. Orange chords represent links identified in this study.

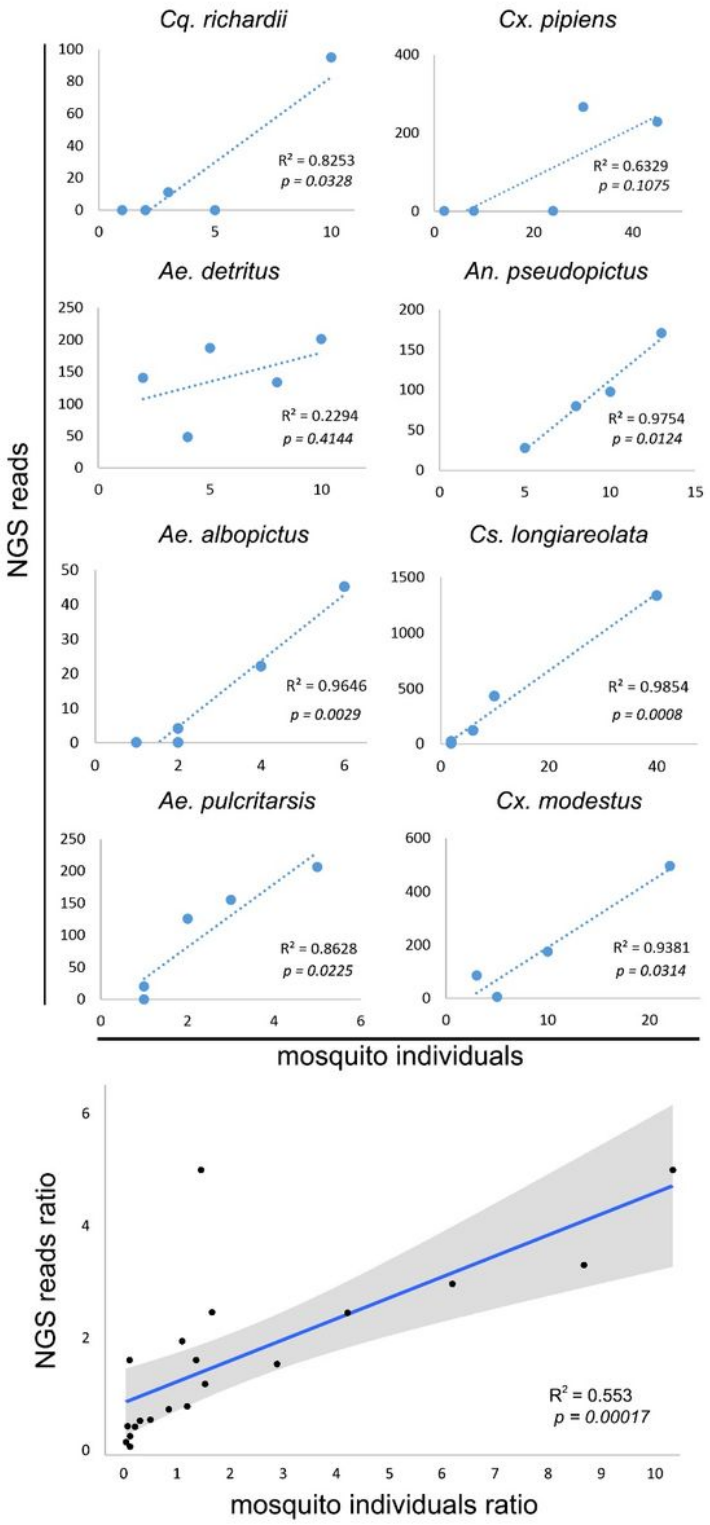

Figure 5

Scatter plots of the quantity of individuals of the different species that were included in the pools and the total number of NGS reads of the respective viruses (upper panel). The ratio of the host/carrier mosquito individuals were plotted against the ratio of the corresponding viruses NGS reads between pools (pool A vs. B, C, D, and E) (lower panel). A linear regression model was applied with the indicated statistics.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.docx