Virology 591 (2024) 109992

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Virome profiling of fig wasps (*Ceratosolen* spp.) reveals virus diversity spanning four realms

Bishwo N. Adhikari [a,b,1], Adrian C. Paskey [a,e,1], Kenneth G. Frey [a], Andrew J. Bennett [a,c,e], Kyle A. Long [a,e], Jens H. Kuhn [d], Theron Hamilton [a], Lindsay Glang [a,e], Regina Z. Cer [a], Tony L. Goldberg [c,f,g], Kimberly A. Bishop-Lilly [a,*]

[a] Genomics and Bioinformatics Department, Biological Defense Research Directorate, Naval Medical Research Command, Frederick, Fort Detrick, MD 21702, USA
[b] Defense Threat Reduction Agency, Fort Belvoir, VA 22060, USA
[c] Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA
[d] Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Fort Detrick, Frederick, MD 21702, USA
[e] Leidos, Inc., Reston, VA 20190, USA
[f] Global Health Institute, University of Wisconsin-Madison, Madison, WI 53706, USA
[g] Department of Zoology, Makerere University, Kampala, Uganda

ARTICLE INFO

Handling Editor: Dr. Jasmine Tomar

Keywords:
Arthropod
Ceratosolen
Ficus
Metagenomics
Fig wasp
Virus discovery

ABSTRACT

We investigated the virome of agaonid fig wasps (*Ceratosolen* spp.) inside syconia ("fruits") of various *Ficus* trees fed upon by frugivores such as pteropodid bats in Sub-Saharan Africa. This virome includes representatives of viral families spanning four realms and includes near-complete genome sequences of three novel viruses and fragments of five additional potentially novel viruses evolutionarily associated with insects, fungi, plants, and vertebrates. Our study provides evidence that frugivorous animals are exposed to a plethora of viruses by coincidental consumption of fig wasps, which are obligate pollinators of figs worldwide.

1. Introduction

Fig wasps (Hymenoptera: Chalcidoidea) form a polyphyletic group that is divided into non-pollinators (several families) and obligate pollinators (family Agaonidae) of fruiting fig trees (Moraceae: *Ficus* spp.). Agaonid wasps of a number of species are highly adapted to fig trees of distinct species, which typically possess both male and female flowers. The co-evolved mutualism of each fig wasp and its fig is in part due to the structure of the fig "fruit," or syconium. Female syconia consist of "inverted flowers", each with an apex that forms a tight ostiole enclosure, providing an exclusive habitat for pollinating wasps. Once the flowers are ready for pollination, the ostiole loosens to enable female wasps to enter, a process during which the wasps lose their wings. After pollinating the flowers, the female wasps die. Male fig wasps are wingless and spend their entire lives within a syconia, where they mate with females and then produce openings through which newly hatched females can escape (Ahmed et al., 2009; Cook et al., 2004; Cook and West, 2005; Weiblen, 2002). Many frugivorous mammals, including pteropodid bats, rely on figs as a food source and therefore ingest fig wasps while feeding (Kunz et al., 1995, 2011). However, with the exception of a tombusvirid in fig wasps (Bennett et al., 2019), the virome of fig wasps and the potential infection of fig-consuming mammals with fig wasp viruses has not yet been examined systematically.

Here, we investigate the viruses of agaonid fig wasps *(Ceratosolen* spp.) sampled in Uganda. *Ceratosolen* fig wasps, which are under taxonomic reassessment, are highly adapted to figs (Jiang et al., 2006), including those that serve as a food source for mammals in Sub-Saharan Africa. Using a metagenomics approach, we provide a glimpse into the fig-wasp-associated virome, which even in this limited investigation reveals viruses spanning four realms. Our results indicate that fig-consuming animals are dietarily exposed to a plethora of thus-far-unknown viruses and that arthropods associated with frugivore

* Corresponding author.
E-mail address: Kimberly.a.bishop-lilly.civ@health.mil (K.A. Bishop-Lilly).
[1] Contributed equally.

https://doi.org/10.1016/j.virol.2024.109992
Received 16 August 2023; Received in revised form 3 January 2024; Accepted 8 January 2024
Available online 10 January 2024

diets may be components of the ecology of mammal-associated viruses.

2. Methods

2.1. Sample collection and homogenizing

Ripe and intact syconia ("fruits") were collected from fig trees of eight species in Kibale National Park (0°13′–0°41″N, 30°19′–30°3″E) and Semliki National Park (0°50′18.5″N and 30°10′22.1″E) Uganda, in January 2016 (Table 1). The fruits were opened using sterile instruments to reveal adult and larval fig wasps inside. Adult fig wasps, both male and female, were collected by tree in pools of 20 (or fewer when less were found). In addition, galls (structures containing larval fig wasps) were collected by tree in pools of 20 (or fewer when less were found). Wasps and larvae were placed into sterile tubes containing 500 μL of DNA/RNA shield buffer (Zymo Research Corporation, Irvine, CA, USA) at a 10:1 buffer:sample ratio. Eleven sample tubes of fig wasps were collected (numbers of insects per tube described in Table 1); one tube (Sample 16) contained only larvae. All samples were exported to the United States with permission of the Uganda Wildlife Authority and the Uganda National Council for Science and Technology and shipped in accordance with International Air Transport Association (IATA) regulations.

2.2. Nucleic acid extraction and sequencing

Nucleic acids were extracted from homogenates using the Qiagen QiAmp MinElute Virus Spin Kit (Qiagen, Germantown, MD, USA) following the manufacturer's protocol. Nucleic acid concentrations were measured using Qubit Broad Range DNA Kit and Qubit High Sensitivity RNA Kit (Thermo Fisher Scientific, Frederick, MD, USA). RNA was converted to double-stranded complementary DNA (cDNA) as previously described (Goldberg et al., 2017). Total cDNA concentration and quantity were assessed using the Qubit Broad Range DNA Kit. Sequencing libraries were prepared with a Nextera XT Kit (Illumina, San Diego, CA, USA) using 1 ng of total DNA as input. Library fragment length and quality were evaluated with a Bioanalyzer 2100 using the High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA, USA). Libraries were normalized and pooled by following manufacturer's protocol and sequenced on the MiSeq platform using MiSeq Reagent Kit v3 (2 × 300 cycles) (Illumina). Selected samples were also processed for additional sequencing with the NEB Next Ultra Directional RNA library kit (New England Biolabs, Ipswich, MA, USA) using the human ribosomal RNA (rRNA) subtraction module following a previously published protocol (Manso et al., 2017). Library fragment length and quality were evaluated as described above and libraries were pooled at 2 nM and sequenced as before, using a spiked-in 1% PhiX control (Illumina, San Diego, CA, USA).

2.3. Virome profiling

Raw reads from all samples were processed using VirusSeeker (Zhao et al., 2017), and resulting data were parsed as previously described (Paskey et al., 2020). In brief, classification of all reads and contigs was performed via the VirusSeeker pipeline using BLASTn, Megablast, and tBLASTx. In-house Python scripts were used to create taxonomic assignment reports for bacteria, fungi, and all other categories that were not generated by VirusSeeker and to calculate the exact read counts corresponding to each contig (VirusSeeker 2.0, manuscript in preparation). The relationship among the number of insects per sample, the total number of reads, and the number of classifiable sequences was investigated by calculating Pearson's correlation coefficient among these variables. R package ggplot2 (Wickham, 2016) was used to create pie charts and heatmaps, and the built-in statistical package R v4.0.3 was used for correlation coefficient analysis (R Core Team, 2018). Raw reads were also submitted to an in-house pipeline called MetaDetector (manuscript in preparation) for initial quality assessment using FASTQC (Andrews, 2014), quality trimming (Q20) using BBDuk (Bushnell, 2014), and assembly by metaSPAdes v3.11.1 (Nurk et al., 2017). Cleaned reads were BLAST-searched using DIAMOND (Buchfink et al., 2015) for taxonomic assignment. MEGAN (Huson et al., 2007) was used to construct rarefaction curves using the read clusters grouped by similarity to estimate species richness at various taxonomic ranks.

2.4. Bioinformatics curation

After classification by VirusSeeker, reads and contigs were binned by established virus families and output as fasta-formatted files. Manual curation of read and/or contig classifications was performed by alignment against the National Center for Biotechnology Information (NCBI) nt or nr database and visual inspection of the results. Misassigned reads and/or contigs were discarded, whereas verified ones were further

Table 1
Summary statistics of metagenomic sequence data from eleven fig wasp samples.

Sample	Number of *Ceratosolen* spp. fig wasps per sample	Presumed species of sampled fig tree	Sampling location	Total sequencing reads	Total base pairs	Number of reads classified as viral	Number of unclassified reads (% of viral reads)
3	20	*Ficus brachylepis* (Welw. ex Hiern)	Kibale National Park	288,562	74,095,627	111	28 (25)
6	20	*Ficus spongii* (?)		574,320	148,825,652	91	43 (47)
7	5	*Ficus brachylepis* (Welw. ex Hiern)		504,982	130,825,014	36	22 (61)
8	10	forest sandpiper figs (*Ficus* (*Urostigma*) *exasperata* Vahl.)		1,731,323	395,113,797	307	113 (37)
11	20	Natal figs (*Ficus* (*Urostigma*) *natalensis* Hochst.)		439,014	117,112,523	289	55 (19)
12	7	false Cape fig (*Ficus* (*Sycomorus*) *vallis-choudae* Delile.)	Semuliki National Park	3,345,203	711,056,399	1760	1250 (71)
13	6	Cape figs (*Ficus* (*Sycomorus*) *sur* Forssk.)		854,697	182,074,908	2898	1265 (44)
14	20	forest sycomore figs (*Ficus* (*Sycomorus*) *mucuso* Welw. ex Ficalho)		672,774	172,891,043	309	105 (34)
15	10	Cape figs (*Ficus* (*Sycomorus*) *sur* Forssk.)		2,002,278	474,619,662	450	122 (27)
16	7	forest sandpiper figs (*Ficus* (*Urostigma*) *exasperata* Vahl.)		2,727,241	637,423,699	145	107 (74)
17	40	Cape figs (*Ficus* (*Sycomorus*) *sur* Forssk.)		2,541,267	588,467,711	109	42 (39)

examined for the presence of conserved viral protein domains. Furthermore, read sets were inspected to ascertain the presence or absence of multiple viral genes and linear breadth of the genome covered by the read set. The purpose of this manual analysis was to ensure that misassigned or false-positive reads and/or contigs be discarded whilst also reporting novel viruses that have no immediate reference and therefore must be compared to their nearest sequenced neighbors. Furthermore, to counter the effect of background contamination that presents a challenge in the analysis of metagenomic sequence data, viral reads detected by VirusSeeker were normalized as Viral Reads Per Million (vRPM) (Toohey-Kurth et al., 2017; Wagner et al., 2012):

$$vRPM = \frac{Number\ of\ reads\ for\ one\ viral\ taxa\ in\ the\ sample}{Total\ number\ of\ QC\ reads\ in\ the\ sample}\ x\ 1,000,000$$

2.5. Viral contig characterization

Contigs and reads that were identified as viral were analyzed using CLC Genomics Workbench v 20.0 (Qiagen, Hilden, Germany). Depth of coverage of each contig was ascertained by stringent read mapping (match score = 1, mismatch cost = 2, similarity fraction = 0.8, length of fraction = 0.8) to the consensus sequence resulting from the assembly (Frey et al., 2014). Protein domain searches were conducted using the Pfam module and database (v32) in CLC Workbench and/or HHpred (Zimmermann et al., 2018). Nucleotide and amino-acid comparisons to the closest sequenced relative were also performed in CLC Workbench. Viruses were assigned to possible host (insect, bacteria, fungus, mammal, plant, or protist) of the most similar virus with a known host where relevant (see "Presumed hosts" column, Supplemental Table 1). Unless specified otherwise, criteria for classification of a virus required that contig(s) shared >90% identity at the amino-acid level and >95% identity at the nucleotide level as compared to the closest relative via the BLAST algorithm. Candidate virus genome sequences that did not share adequate identity with their closest sequenced neighbors were named and numbered "wugcerasp [family name]-like virus [number]. This name is derived from the geography and the host associated with the new viruses, taking "wug" for Western Uganda and "cerasp" for *Ceratosolen* spp. wasp. Each novel contig received a separate sequential number, even when the contig shared the same nearest sequenced neighbor because we cannot assume that each contig belongs to the same novel virus in absence of knowing the complete genome sequence. In case future studies reveal that these multiple contigs represent the same virus, numbers can be condensed into one to avoid the possibility of double naming/numbering in the future.

2.6. Phylogeny reconstruction

Open reading frames (ORFs) that encode viral RNA-dependent RNA polymerase (RdRp) genes were used for phylogenetic analyses. The large (L) protein genes of the large (L) genomic segments, which encode RdRp domains, were used for bunyavirals (phasmavirids and phenuivirids). Nucleotide sequences representing the ORFs of viral contigs and near-neighbors were aligned using CLC Workbench v23 (QIAGEN) and maximum-likelihood trees were generated using IQ-TREE2 (Minh et al., 2020), ModelFinder (Kalyaanamoorthy et al., 2017), tree reconstruction, and ultrafast bootstrap (1000 replicates) (Hoang et al., 2018). Trees were visualized using FigTree (Rambaut, 2009) and each rooted at the midpoint. Models selected via ModelFinder were used for maximum-likelihood trees are as follows: TIM3+F + G4 for tombus-like viruses (log likelihood −38,263, Fig. 3A); GTR + F + R3 for fig wasp orthophasmavirus (log likelihood −102,368, Fig. 3B) and fig wasp iflaviruses 1 and 2 (log likelihood −197,966, Fig. 3D); GTR + F + R4 for fig wasp goukovirus (log likelihood −139,772, Fig. 3C) and fig wasp partiti-like virus (log likelihood −41,404, Fig. 3E); and TVMe + G4 for fig wasp narna-like virus and fig wasp Humaitá-Tubiacanga virus (log likelihood −22,142, Fig. 3F).

3. Results

3.1. Preliminary metagenomic analyses

Pooled, homogenized fig wasps were processed for sequencing, and libraries were prepared such that both DNA extract and cDNA generated from the RNA extract were included as input. All but sample 16 contained adult wasps. The average number of reads per sample was ≈1.5 million (Table 1). As might be expected, due to normalization of libraries prior to sequencing, there was no correlation between the number of insects per sample and the subsequent number of generated sequence reads ($r = -0.05$ and $p < 0.001$). Similarly, there was no correlation between the number of sequence reads and the number of classifiable sequences ($r = 0.166$ and $p < 0.01$). Furthermore, we found no correlation between the number of fig wasps per pool and the number of classifiable sequences ($r = -0.041$ and $p < 0.01$). Reads were predominantly binned by VirusSeeker as "other" (not classified) or bacteria. Samples 12 and 13 had the greatest proportion of viral reads (Fig. 1). The virome content of most samples consisted primarily of "insect or fungus associated viruses" followed by "unclassified RNA/DNA/environmental viruses" as determined using VirusSeeker.

To perform an initial assessment of the overall genetic richness contained within the samples, we constructed rarefaction plots using read clusters grouped by similarity. Despite the differences in the number of fig wasps in the samples, the rarefaction curves reached their asymptotes or started to plateau for all taxa ranks in nearly all samples, indicating that read depth was most likely sufficient for all samples (Fig. S1). The asymptotes were lowest for samples 6 and 8, indicating the lowest level of genetic richness. The asymptotes for samples 3, 11, 12, 13, 14, and 15 indicated an intermediate level of genetic richness. The asymptotes were comparatively higher for samples 7, 16, and 17, in which genetic richness was highest. These rarefaction asymptotes imply that further sequencing would likely not have resulted in identification of additional viral contigs in most fig wasp samples.

3.2. Virome profile

Taxonomic profiling of the fig wasp (*Ceratosolen* spp.) virome resulted in the identification of viruses belonging to 28 families of four realms (*Duplodnavirida*, *Monodnaviria*, *Riboviria*, and *Varidnaviria*) (Koonin et al., 2020) and one viriform (Kuhn and Koonin, 2023) from diverse hosts. Supplemental Table 1 documents the relative abundance of each detected viral family. Compared to other families, a higher proportion of unclassified reads matched ribovirian families *Peribunyaviridae*, *Orthomyxoviridae*, and *Rhabdoviridae*, all of which include insect viruses. As expected, based on sample type (i.e., fig wasps), the majority of detected virus and viriform nucleic acids are related to known insect-related viruses; but profiling also identified reads classified as viruses associated with fungi, mammals, plants, and protists. In addition to the overall taxonomic diversity among the samples, there was noticeable diversity among sequences from individual viruses, suggesting distinct virus variants infecting individual fig wasps within a given sample. Samples consisting of ≥10 fig wasps contained higher numbers of viruses (≥4) than samples containing <10 fig wasps. More than a third of the contigs that could be assigned were related to known viruses with positive-sense RNA genomes, followed by contigs matching double-stranded DNA virus genomes. Contigs related to negative-sense RdRp-encoding RNA viruses (ribovirian phylum *Negarnaviricota*), which were mostly insect-related, matched orders *Bunyavirales* (families *Nyamiviridae*, *Peribunyaviridae*, and *Rhabdoviridae*) and *Articulavirales* (family *Orthomyxoviridae* sensu lato). Notably, unclassified bunyavirals were substantially represented in the samples, comprising over 25% of the hits in fig wasp samples 12, 13, and 17. Additionally, unclassified rhabdovirids (order *Mononegavirales*) made up five contigs and over 10%

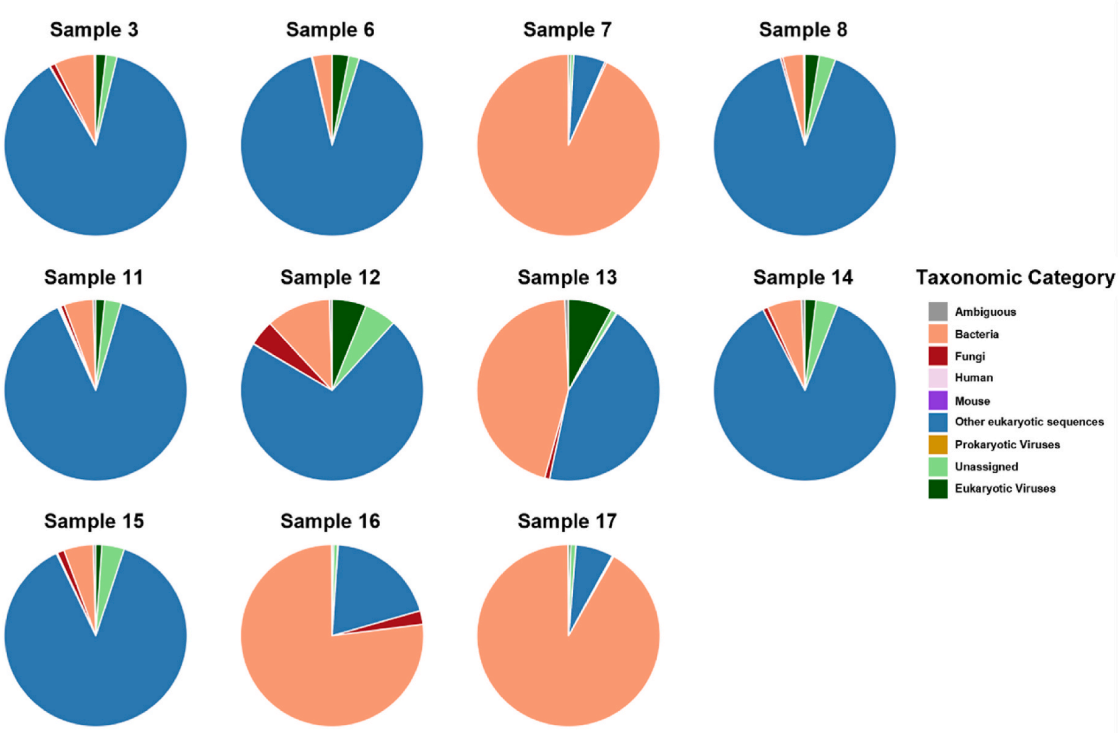

Fig. 1. Microbial profiling of fig wasp (*Ceratosolen* spp.) samples. Pie charts show the proportion of metagenomic reads assigned to different taxonomic categories in samples. Taxonomic assignment is based on classification by VirusSeeker.

of all viral reads in sample 6. The metagenome assembly derived from sample 12 also contained five contigs corresponding to unclassified rhabdovirids, as well as four reads corresponding to an unclassified quaranjavirus (*Orthomyxoviridae*). With the exception of samples that resulted in overall few sequencing reads and relatively fewer unclassified viruses, other virus families of interest represented low normalized abundance (Fig. 2).

3.3. Characterization of novel viruses

Overall, a large portion of the virus-derived reads (up to 74%) obtained from the fig wasp samples could not be matched to classified viruses, suggesting that these reads were derived from novel viruses. Based on overall deduced amino-acid sequence identity and genome structure and organization, these assembled reads could, however, be

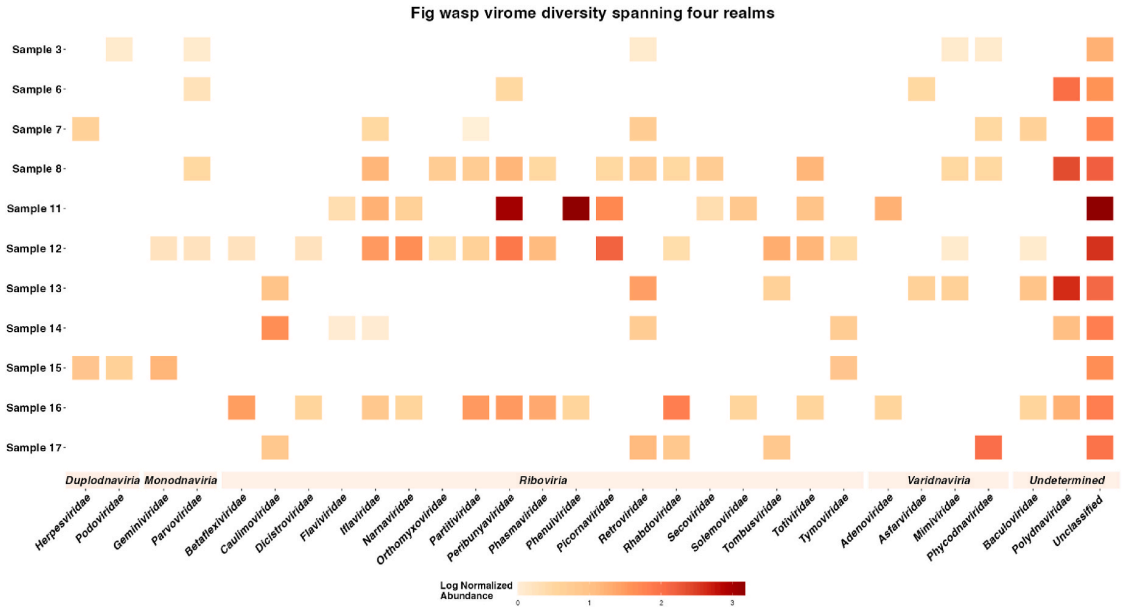

Fig. 2. Fig wasp (*Ceratosolen* spp.) virome diversity spanning four realms. The heat map indicates log-normalized abundance of metagenomic reads assigned to established virus families. Virus families are grouped along the x-axis by virus realm (*Duplodnaviridae*, *Monodnaviria*, *Riboviria*, *Varidnaviria*, and undetermined).

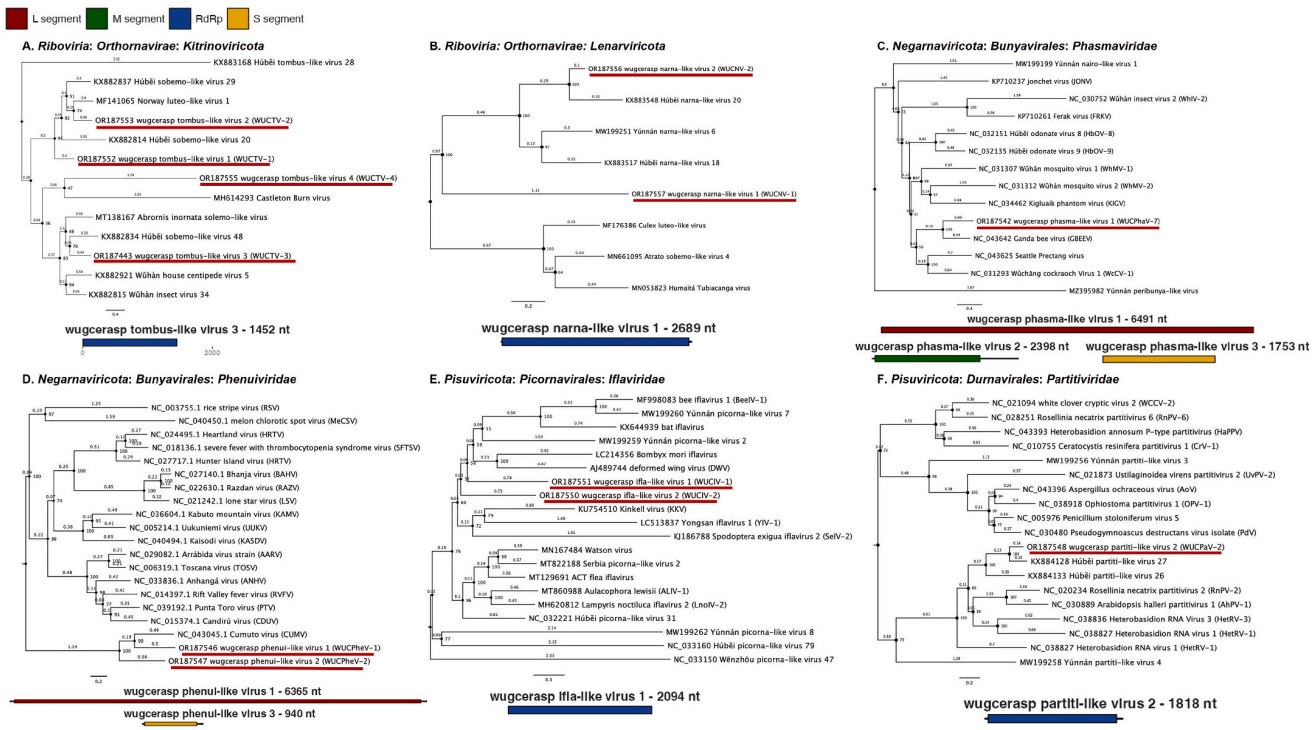

Fig. 3. Phylogenetic analysis of partial genomes of putatively novel viruses identified in fig wasps (*Ceratosolen* spp.). Phylogenetic analysis performed using alignment of known and newly identified RNA-dependent RNA polymerase (RdRp) nucleotide sequences using the maximum-likelihood method. Branch lengths and bootstrap values are included for each tree (1000 replicates). Novel viruses are underlined in red and genome structures representing the new virus fragments are shown as part of each subfigure (legend top left: red for L segments, green for M segments, gold for S segments, and blue for RdRp fragments).

assigned to established families (Fig. 2). Datasets derived from samples 12 and 13 included the largest proportions (71 and 44%, respectively) of novel virus sequences (Fig. 2). Datasets derived from samples 6, 14, 15, and 16 also contained an appreciable number of unclassified virus sequences. Using the sequence reads from these datasets, we were able to determine near-complete genome sequences of three novel viruses and fragments of five additional novel virus genomes.

3.4. *Wugcerasp tombus-like virus 1 and 2 are novel tombusvirids (Tolivirales)*

From samples 6 and 12, we assembled 1.3-kb and 0.8-kb contigs, both corresponding to the RdRp ORF of a novel tombusvirid (ribovirian order *Kitrinoviricota*, order *Tolivirales*). The proteins encoded by these contigs share 53.9 and 62.1% similarity at the amino-acid level with RdRp of Norway luteo-like virus 1, an unclassified tombusvirid discovered in castor bean ticks (*Ixodes ricinus* (Linnaeus, 1758)) (Pettersson et al., 2017) (Table 2). Phylogenetic analysis revealed that these two fig wasp viruses fall within different subclades (Fig. 3A). To our knowledge, this is the first report of tombusivirids in fig wasps. We named the novel viruses wugcerasp tombus-like virus 1 (WUCTV-1) and wugcerasp tombus-like virus 2 (WUCTV-2). Based on the current demarcation criteria considering the number of genome segments and size of the genome, as well as polymerase characterization as described by the International Committee on the Taxonomy of Viruses (ICTV) *Tombusviridae* Study Group (King et al., 2011), it remains unclear whether these two new viruses belong to an established genus.

Other contigs that may indicate the presence of novel tombusvirids: Another RdRp sequence is 53.1% similar to the RdRp of Wǔhàn house centipede virus 5 and we have called it wugcerasp tombus-like virus 3 (WUCTV-3). Phylogenetically, this virus clusters with Húběi sobemo-like virus 48 and related viruses. Another RdRp sequence is moderately similar (58.7%) to the RdRp of Castleton Burn virus (CBV) sequenced from bumblebees (apid *Bombus* sp.) (Pascall et al., 2021) (Fig. 3A) and only 46% similar to the RdRp of Húběi tombus-like virus 28 discovered in long-jawed orb-weavers (tetragnathid *Tetragnatha keyserlingi* Simon, 1890; Shi et al., 2016). Phylogenetic analysis places the fig-wasp-associated sequence, called wugcerasp tombus-like virus 4 (WUCTV-4), near CBV.

3.5. *Wugcerasp narna-like virus 1 and 2 are novel wolframviruses (Orthornavirae: Lenarviricota)*

We used RdRp sequence comparisons to identify additional potentially novel viruses in the fig wasp samples (Fig. 3). The polypeptide encoded by one of the contigs (2.7 kb) was moderately similar (57.1%) to the RdRp of Humaitá-Tubiacanga virus and Culex mosquito virus 6 previously found in yellow fever mosquitoes (culicid *Aedes aegypti* (Linnaeus in Hasselquist, 1762)) (Aguiar et al., 2015; Parry and Asgari, 2018; Shi et al., 2019) and other arthropods (Webster et al., 2016). Phylogenetic analysis of the fig-wasp-associated sequence, called wugcerasp narna-like virus 1 (WUCNV-1), suggests the existence of a fig wasp virus distantly related to narnavirids (ribovirian *Lenarviricota*, order *Wolframavirales*) and possibly representing a new wolframviral family (Fig. 3B).

The last RdRp sequence is wugcerasp narna-like virus 2 (WUCNV-2), which shares relatively highly similar (72.7%) to the RdRp of Húběi narna-like virus 20 discovered in dipteran insects (Shi et al., 2016) and clusters with this sequence as shown in Fig. 3B.

3.6. *Wugcerasp phasma-like virus 1, 2, and 3 contigs comprise segments of a novel orthophasmavirus (Bunyavirales: Phasmaviridae)*

We obtained a near-complete genome sequence of a potentially novel

Table 2
Viral sequences that may represent novel viruses.

Virus name	NCBI accession	Genome type	Length of putative viral contig (nt)	BlastX result	GenBank accession number	Amino-acid similarity (%)	Query coverage (%)	Deduced family
Riboviria: *Orthornavirae*: ***Kitrinoviricota***								
wugcerasp tombus-like virus 1	OR187552	Unsegmented positive-sense RNA	1283	Norway luteo-like virus 1 RdRp	ASY03252	53.86	99	*Tombusviridae*
wugcerasp tombus-like virus 2	OR187553	Unsegmented positive-sense RNA	807	Norway luteo-like virus 1 RdRp	ASY03252	62.13	99	*Tombusviridae*
wugcerasp tombus-like virus 3	OR187554	positive-sense RNA	1452	Wǔhàn house centipede virus 5 RdRp	YP_009342464	53.16	97	*Tombusviridae*
wugcerasp tombus-like virus 4	OR187555	positive-sense RNA	1141	Castleton Burn virus RdRp	QAY29244	58.68	99	*Tombusviridae*
Riboviria: *Orthornavirae*: ***Lenarviricota***								
wugcerasp narna-like virus 1	OR187557	positive-sense RNA	2689	Humaitá-Tubiacanga virus RdRp	QEM39312	57.11	54	*Narnaviridae*
wugcerasp narna-like virus 2	OR187556	unsegmented positive-sense RNA	677	Húběi narna-like virus 20 RdRp	APG77159	72.65	99	*Narnaviridae*
Riboviria: *Orthornavirae*: ***Negarnaviricota***								
wugcerasp phasma-like virus 1	OR187542	trisegmented negative-sense RNA	6491	Ganda bee virus RdRp gene (L segment)	YP_009666981	55.6	94	*Phasmaviridae*
wugcerasp phasma-like virus 2	OR187544		2398	Ganda bee virus glycoprotein precursor (M segment)	YP_009666982	41.28	69	*Phasmaviridae*
wugcerasp phasma-like virus 3	OR187543		1753	Ganda bee virus nucleoprotein (S segment)	YP_009666983	47.23	52	*Phasmaviridae*
wugcerasp phenui-like virus 1	OR187546	trisegmented negative-sense RNA	6365	Cumuto virus RdRp (L segment)	YP_009664615	50.98	95	*Phenuiviridae*
wugcerasp phenui-like virus 2	OR187547		6267	Cumuto virus RdRp (L segment)	YP_009664615	46.72	97	*Phenuiviridae*
wugcerasp phenui-like virus 3	OR187545		940	Cumuto virus nucleocapsid (S segment)	YP_009664617	42.91	82	*Phenuiviridae*
Riboviria: *Orthornavirae*: ***Pisuviricota***								
wugcerasp ifla-like virus 1	OR187551	unsegmented positive-sense RNA	2094	Aulacophora lewisii iflavirus 1 polyprotein	QOW95919	42.54	82	*Iflaviridae*
wugcerasp ifla-like virus 2	OR187550		1918	ACT flea iflavirus polyprotein	QIJ70026	40.82	91	*Iflaviridae*
wugcerasp ifla-like virus 3	–		1864	Aulacophora lewisii iflavirus 1 polyprotein	QOW95919	42.83	84	*Iflaviridae*
wugcerasp partiti-like virus 1	OR187548	bipartite (RdRp and CP) double-stranded RNA	4208	Húběi partiti-like virus 27 RdRp	APG78241	82.37	76	*Partitiviridae*
wugcerasp partiti-like virus 2	OR187549		1,818	Húběi partiti-like virus 27 RdRp	APG78241	78.72	91	*Partitiviridae*
wugcerasp solemo-like virus	OR187541	unsegmented positive-sense RNA	817	Neohydatothrip associated sobemo-like virus 1 putative CP	QNM37826	32.72	98	*Solemoviridae*

CP, capsid protein; RdRp, RNA-directed RNA polymerase.

orthophasmavirus. Orthophasmaviruses have trisegmented negative-sense RNA genomes consisting of L, medium (M), and small (S) segments. We identified a 6.5-kb contig with 67.4% nucleotide similarity to the L-segment sequence of a Ganda bee virus (GBEEV; species *Orthophasmavirus gandaense*), a virus known to infect western honey bees (apid *Apis mellifera* Linnaeus, 1758) (Schoonvaere et al., 2016) (Table 2). The encoded L protein amino-acid sequence, including the RdRp domain, is 55.6% similar to that of GBEEV and also included a conserved region from the general bunyaviral RdRp. The M and S segment-encoded protein sequences of the new virus are 41.3 and 47.2% similar to the GBEEV glycoprotein precursor and nucleoprotein (N) sequences, respectively. The glycoprotein sequence was found to contain a conserved glycoprotein domain typical for phasmavirids. Samples 7 and 13 contained 3072–8541 (141–277× coverage) reads mapped to the three contigs of the novel virus (Fig. S2). Phylogenetic reconstruction using the putative RdRp sequence and phasmavirid RefSeq genomes indicates clustering of the novel virus with GBEEV (Fig. 3C). The new virus represents a new orthophasmavirus species according to the current species demarcation criteria defined by the ICTV *Phasmaviridae* Study Group. (Species demarcation criteria for the family *Phasmaviridae* are <95% identity in the amino-acid sequence of the entire RdRp.) We named the novel fig wasp orthophasmavirus contigs wugcerasp phasma-like virus 1 (WUCPhaV-1), wugcerasp phasma-like virus 2 (WUCPhaV-2), and wugcerasp phasma-like virus 3 (WUCPhaV-3).

3.7. Wugcerasp phenui-like virus 1, 2, 3 contigs represent novel goukoviruses (Bunyavirales: Phenuiviridae)

From fig wasp samples 3 and 12, we determined a near-complete genome sequence of a potentially new goukovirus. Goukoviruses typically have trisegmented negative-sense RNA genomes consisting of L, M, and S segments encoding an L protein, a glycoprotein precursor, and N protein, respectively. We identified 6.3-kb and 6.4-kb contigs encoding proteins with 46.7 and 51% amino-acid sequence similarities to the L protein of Cumuto virus (CUMV; species *Goukovirus cumutoense*), a virus known to infect mosquitoes (culicid *Culex declarator* Dyar and Knab, 1906) (Auguste et al., 2014) (Table 2). Both contigs also encode a conserved region from the general bunyaviral RdRp. In addition, we found a nearly 1-kb contig encoding a protein with 42.9% sequence similarity to CUMV N, as well as a conserved phlebovirus/tenuivirus N domain. Further mapping of the reads to novel viral segments showed 10,736 and 1265 reads mapped to the corresponding L and S segments with median coverage of 361× and 270×, respectively (Fig. S3). Contigs corresponding to M segment/glycoprotein sequences were not detected. Given the relatively deep coverage associated with the L and S segments (361× and 270×, respectively), we hypothesize that this apparent lack of an M segment was not due to insufficient sampling, but indeed is not needed by this novel virus. Phylogenetic reconstruction using the putative RdRp sequence and phenuivirid RefSeq genomes indicates clustering of the novel virus with CUMV (Fig. 3D). The new virus represents a new goukovirus species according to the current species demarcation criteria defined by the ICTV *Phenuiviridae* Study Group by applying the established demarcation criterion of <95% amino-acid identity among RdRp domains. We named the novel virus fig wasp goukovirus (*Phenuiviridae*) contigs wugcerasp phenui-like virus 1 (WUCPheV-1), wugcerasp phenui-like virus 2 (WUCPheV-2), and wugcerasp phenui-like virus 3 (WUCPheV-3).

3.8. Wugcerasp ifla-like virus 1, 2, and 3 are novel iflavirids (Picornavirales)

In samples 6, 12, and 13, we identified three novel contigs (2.1, 1.9, and 19 kb, respectively) that are highly similar to genomes of iflavirids. Specifically, these contigs matched iflavirus polyprotein regions of ACT flea iflavirus of rabbit fleas (pulicid *Spilopsyllus cuniculi* (Dale, 1878)) and Aulacophora lewisii iflavirus 1 of pumpkin beetles (chrysomelid *Aulacophora lewisii* Baly, 1886; Ye et al., 2021) (Table 2). The longest contig encodes the polyprotein gene, including two conserved picornaviral capsid protein domains. The second contig encodes a conserved region of the picornaviral RdRp, whereas the third contig encodes a picornaviral capsid protein domain. The deduced amino-acid sequence similarities of these contigs to were low enough (below 43%) to ACT flea iflavirus Aulacophora lewisii iflavirus 1 to suggest that they represent several new viruses. Phylogenetic analysis of the fig wasp-associated RdRp sequences resulted in clustering of fig wasp iflavirus 1 with Húběi picorna-like virus 79 and Wēnzhōu picorna-like virus 47. The contigs representing fig wasp iflaviruses 1 and 2 fall within the same clade, although they are distinct viruses (Fig. 3E). Based on the current species demarcation criteria established by the ICTV *Iflaviridae* Study Group (≤90% amino-acid sequence similarity of capsid protein) (Valles et al., 2017), they were called wugcerasp ifla-like virus 1 (WUCIV-1), wugcerasp ifla-like virus 2 (WUCIV-2), and wugcerasp ifla-like virus 3 (WUCIV-3).

3.9. Wugcerasp partiti-like virus 1 and 2 represent a novel alphapartitivirus (Durnavirales: Partitiviridae)

A near-complete genome sequence of a potentially new alphapartitivirus was assembled from samples 7 and 13. Alphapartitiviruses have bipartite double-stranded RNA genomes. We identified two contigs of 4.2 and 1.8 kb encoding proteins with 82.4 and 78.7% sequence similarity to the RdRp encoded by Segment 1 of Húběi partiti-like virus 27, a virus that was previously identified via sequencing from odonatan insects (Shi et al., 2016) (Table 2). Further mapping of the reads to novel viral segments demonstrated that 43,320 and 43,324 reads mapped to the corresponding Segment 1 with median coverage of 2,432× and 4,671×, respectively (Fig. S4). Phylogenetic analysis resulted in clustering of the new virus sequence with a monophyletic clade, including Húběi partiti-like viruses 26 and 27, and showed that partitivirids associated with fungi and plants form distinct clades (Fig. 3F). Based on the species demarcation threshold in the *Alphapartitivirus* genus proposed by the ICTV (≤90% amino-acid sequence similarity in the RdRp and ≤80% amino-acid sequence similarity in the capsid protein encoded by genome Segment 2 (Vainio et al., 2018)), the fig wasp virus may represent a new alphapartitivirus species and is likely the first reported partitivirid from fig wasps.

4. Discussion

Arthropods harbor a range of viruses that are often divergent from taxonomic relatives infecting vertebrates (Edgar et al., 2022; Käfer et al., 2019; Li et al., 2015; Neri et al., 2022; Paraskevopoulou et al., 2021; Shi et al., 2016). Over the past decade, advances in high-throughput sequencing and metagenomics have greatly expanded our knowledge of the plethora of viruses in nature (Callanan et al., 2020; Edgar et al., 2022; Gregory et al., 2019; Hou et al., 2023; Käfer et al., 2019; Lee et al., 2023; Li et al., 2015; Neri et al., 2022; Olendraite et al., 2023; Paraskevopoulou et al., 2021; Shi et al., 2018; Shi et al., 2016; Tisza et al., 2020; Zayed et al., 2022), including arthropod-borne viruses associated with diseases of vertebrates (Bolling et al., 2015; Vasilakis and Tesh, 2015). We have hypothesized that viruses hosted by invertebrates that are intimately associated with mammals in nature, such as biting arthropods (e.g., bat flies) or arthropods inhabiting food items that, for instance, bats consume, could be of epidemiological relevance (Bennett et al., 2019, 2020; Blomström et al., 2020; Goldberg et al., 2017; Jansen van Vuren et al., 2017; Kamani et al., 2022; Kuang et al., 2023; Ortiz-Baez et al., 2023; Peng et al., 2022; Ramírez-Martínez et al., 2021; Xu et al., 2022). Frequent exposure of mammals to such viruses could lead to occasional and/or transient mammal infections, perhaps in some cases leading to the evolution of novel mammal-borne viruses. Such interactions could also explain the detection of fragments or tangential evidence of non-mammal viruses in mammals using surveillance tools such as PCR or serology. We note that it is possible that the fig wasp viruses identified in this study could originate from the fig wasp diet or from fig wasp parasites.

Fig wasps are obligate endosymbionts of figs, a major food source for pteropodid bats in sub-Saharan Africa (Fleming and Kress, 2011). Many frugivorous mammals, including bats, rely on figs as a food source and therefore ingest fig wasps while consuming figs (Kunz et al., 1995, 2011). Our results revealed a remarkable diversity of viruses with similarities to insect-, plant-, and mammal-associated viruses. We identified viruses from 28 families, spanning four realms, including viruses with diverse host ranges. Among them were several unclassified viruses belonging to established families and novel viruses related to those hosted by fungi, insects, plants, and vertebrates. This diversity could partly be due to the roles of viruses in almost all ecological guilds, where they interact with eukaryotes other than figs, including other plants and animals (Segura-Trujillo et al., 2016). The discoveries here point toward a broad diversity of likely fig wasp-specific viruses and interactions. As shown in Fig. 2, the fig wasps studied here contain viruses covering the diversity of nearly the entire virosphere (Koonin et al., 2020). A recent extensive study of bat parasites identified overall similar viruses within the order *Bunyavirales*, as well as pariti-like, picorna-like and narna-like viruses (Fig. 3) (Xu et al., 2022). The global diversity of fig wasp viruses is likely to be vastly higher than that reported herein, given the extraordinary diversity and abundance of figs worldwide (Harrison, 2005).

One of the key findings of our study is that much of the virus diversity in these fig wasps lies within the order *Bunyavirales*, which includes viruses from arthropods and mammals that can cause serious human diseases (Barr et al., 2020). For instance, we discovered a novel orthophasmavirus belonging to family *Phasmaviridae*, a family for viruses that are commonly associated with diverse insect hosts, ranging from agricultural pests to vectors of human disease (Shi et al., 2016), and a novel goukovirus belonging to the large family *Phenuiviridae*. Whether the examined fig wasps are the exclusive hosts of these two viruses or whether they are able to infect other arthropods, such as bees, or even animals that feed on fruit and wasps remains to be examined.

We also identified fig-wasp-associated viruses that belong to the picornaviral family *Iflaviridae,* which was originally established to include diverse insect-related viruses; recent metagenomic analyses have expanded the known host range of these viruses to include both bats and arthropods (Sakuna et al., 2017; Shi et al., 2016; Yinda et al., 2017). Three of the novel viruses identified in this study clustered with insect iflavirids. The clustering of these viruses within a clade rich in arthropod-associated viruses suggests that the novel iflavirids are possibly restricted to fig wasps specifically or arthropods in general. Conversely, viruses of family *Partitiviridae* are known to be mostly associated with fungi and plant hosts (Vainio et al., 2018). As such, identification of a novel partitivirid that clusters with arthropod-associated alphapartitiviruses suggests that these viruses could plausibly have arisen from a fungus in or on the fig wasp, in or on the fig fruit, or the tree itself.

Finally, we found a novel luteovirus (*Tombusviridae*). This genus includes many economically important plant pathogens (Ali et al., 2014; Gray and Gildow, 2003), but wugcerasp tombus-like viruses 1 and 2 cluster with tick-associated luteoviruses that are distinct from luteo-like viruses identified from other arthropods.

Similar to other arthropod virome studies, a considerable portion of determined viral sequences in this study are unassignable to established taxa (Shi et al., 2016), emphasizing that the arthropod virosphere in general, and the fig wasp virome in particular, is far from being delineated.

5. Conclusion

Metagenomic shotgun sequencing data produced in this study demonstrate that wasps that feed on the fruits of ficus trees in Sub-Saharan Africa contain a diversity of viruses, including viruses that are novel or previously uncharacterized, as well as viruses that are currently unassignable to established taxa. As these trees are also frequented by frugiverous mammals that may ingest these fruits and the associated wasps, this represents a potential source of mammal exposure to diverse viruses. Therefore, the arthropod virosphere deserves further study and characterization to improve our knowledge of the viruses to which mammals are exposed in their various natural habitats.

Funding information

This study was supported by Armed Forces Health Surveillance Division (AFHSD) Global Emerging Infections Surveillance (GEIS) Branch through P0032_18_NM_05, P0071_19_NM_06, P0062_20_NM_06, and P0094_21_NM to K.A.B. This work was supported in part through the Laulima Government Solutions, LLC, prime contract with the U.S. National Institute of Allergy and Infectious Diseases (NIAID) under Contract No. HHSN272201800013C. J.H.K. performed this work as an employee of Tunnell Government Services (TGS), a subcontractor of Laulima Government Solutions, LLC, under Contract No. HHSN272201800013C.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Health and Human Services, the U.S. Department of Defense, the U.S. Navy, the U.S. government or of the institutions and companies affiliated with the authors.

Several of the authors are U.S. government employees. This work was prepared as part of their official duties. Title 17 U.S.C. § 105 states, "Copyright protection under this title is not available for any work of the United States Government." Title 17 U.S.C. §101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person's official duties.

Ethical approval

No human or animal approvals were required for this work.

Repositories

Trimmed reads for the 11 samples are available under NCBI Project Number PRJNA986526 in Sequence Read Archive (SRA). The 18 new virus contigs are available in GenBank under accession numbers OR187541–OR187557 and OR269138.

Impact statement

We aimed to understand which viruses infect a particular category of wasps that feed on the fruits of ficus trees also frequented by fruit-eating mammals in Sub-Saharan Africa. We discovered a broad diversity of novel wasp viruses that are related to viruses of four of the major branches of the virosphere. Thus, we show that fig-eating mammals, via eating wasp-infested figs, are exposed to these viruses.

Data summary

All data generated through this work have been deposited in NCBI SRA (host-removed, trimmed shotgun reads for 11 samples; Project PRJNA986526) and GenBank (18 contigs representing novel viruses; accession numbers OR187541–OR187557 and OR269138). Supplementary materials include richness estimates for each sample, as well as read mappings for relevant contigs of novel viruses.

CRediT authorship contribution statement

Bishwo N. Adhikari: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Adrian C. Paskey:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Kenneth G. Frey:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Andrew J. Bennett:** Investigation, Formal analysis. **Kyle A. Long:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Jens H. Kuhn:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Theron Hamilton:** Investigation, Conceptualization. **Lindsay Glang:** Formal analysis. **Regina Z. Cer:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Formal analysis. **Tony L. Goldberg:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Kimberly A. Bishop-Lilly:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Anya Crane (National Institutes of Health [NIH]

National Institute of Allergy and Infectious Diseases [NIAID] Division of Clinical Research [DCR] Integrated Research Facility at Fort Detrick [IRF-Frederick], Frederick, MD, USA) for editing the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2024.109992.

References

Aguiar, E.R., Olmo, R.P., Paro, S., Ferreira, F.V., de Faria, I.J., Todjro, Y.M., Lobo, F.P., Kroon, E.G., Meignin, C., Gatherer, D., Imler, J.-L., Marques, J.T., 2015. Sequence-independent characterization of viruses based on the pattern of viral small RNAs produced by the host. Nucleic Acids Res. 43, 6191–6206.

Ahmed, S., Compton, S.G., Butlin, R.K., Gilmartin, P.M., 2009. Wind-borne insects mediate directional pollen transfer between desert fig trees 160 kilometers apart. Proc. Natl. Acad. Sci. U.S.A. 106, 20342–20347.

Ali, M., Hameed, S., Tahir, M., 2014. *Luteovirus*: insights into pathogenicity. Arch. Virol. 159, 2853–2860.

Andrews, S., 2014. FastQC: a quality-control tool for high-throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Auguste, A.J., Carrington, C.V.F., Forrester, N.L., Popov, V.L., Guzman, H., Widen, S.G., Wood, T.G., Weaver, S.C., Tesh, R.B., 2014. Characterization of a novel *Negevirus* and a novel *Bunyavirus* isolated from *Culex (Culex)* declarator mosquitoes in Trinidad. J. Gen. Virol. 95, 481–485.

Barr, J.N., Weber, F., Schmaljohn, C.S., 2020. *Bunyavirales*: the viruses and their replication. In: Howley, P.M., Knipe, D.M., Whelan, S.P.J. (Eds.), Fields Virology, seventh ed. Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia, Pennsylvania, USA, pp. 706–749.

Bennett, A.J., Bushmaker, T., Cameron, K., Ondzie, A., Niama, F.R., Parra, H.-J., Mombouli, J.-V., Olson, S.H., Munster, V.J., Goldberg, T.L., 2019. Diverse RNA viruses of arthropod origin in the blood of fruit bats suggest a link between bat and arthropod viromes. Virology 528, 64–72.

Bennett, A.J., Paskey, A.C., Kuhn, J.H., Bishop-Lilly, K.A., Goldberg, T.L., 2020. Diversity, transmission, and cophylogeny of ledanteviruses (*Rhabdoviridae: Ledantevirus*) and nycteribiid bat flies parasitizing Angolan soft-furred fruit bats in Bundibugyo District, Uganda. Microorganisms 8, 750.

Blomström, A.-L., Luz, H.R., Öhlund, P., Lukenge, M., Brandão, P.E., Labruna, M.B., Berg, M., 2020. Novel viruses found in *Antricola* ticks collected in bat caves in the western Amazonia of Brazil. Viruses 12, 48.

Bolling, B.G., Weaver, S.C., Tesh, R.B., Vasilakis, N., 2015. Insect-specific virus discovery: significance for the arbovirus community. Viruses 7, 4911–4928.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

Bushnell, B., 2014. BBMap: a fast, accurate, splice-aware aligner. LBNL Report Number LBNL-7065E. Lawrence Berkley National Laboratory. https://www.osti.gov/biblio/1241166-bbmap-fast-accurate-splice-aware-aligner.

Callanan, J., Stockdale, S.R., Shkoporov, A., Draper, L.A., Ross, R.P., Hill, C., 2020. Expansion of known ssRNA phage genomes: from tens to over a thousand. Sci. Adv. 6, eaay5981.

Cook, J.M., Bean, D., Power, S.A., Dixon, D.J., 2004. Evolution of a complex coevolved trait: active pollination in a genus of fig wasps. J. Evol. Biol. 17, 238–246.

Cook, J.M., West, S.A., 2005. Figs and fig wasps. Curr. Biol. 15, R978–R980.

Edgar, R.C., Taylor, J., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., Banfield, J.F., de la Peña, M., Korobeynikov, A., Chikhi, R., Babaian, A., 2022. Petabase-scale sequence alignment catalyses viral discovery. Nature 602, 142–147.

Fleming, T.H., Kress, W.J., 2011. A brief history of fruits and frugivores. Acta Oecol. 37, 521–530.

Frey, K.G., Herrera-Galeano, J.E., Redden, C.L., Luu, T.V., Servetas, S.L., Mateczun, A.J., Mokashi, V.P., Bishop-Lilly, K.A., 2014. Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. BMC Genom. 15, 96.

Goldberg, T.L., Bennett, A.J., Kityo, R., Kuhn, J.H., Chapman, C.A., 2017. Kanyawara virus: a novel rhabdovirus infecting newly discovered nycteribiid bat flies infesting previously unknown pteropodid bats in Uganda. Sci. Rep. 7, 5287.

Gray, S., Gildow, F.E., 2003. Luteovirus-aphid interactions. Annu. Rev. Phytopathol. 41, 539–566.

Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Cruaud, C., Dimier, C., Domínguez-Huerta, G., Ferland, J., Kandels, S., Liu, Y., Marec, C., Pesant, S., Picheral, M., Pisarev, S., Poulain, J., Tremblay, J.-É., Vik, D., Babin, M., Bowler, C., Culley, A.I., de Vargas, C., Dutilh, B.E., Iudicone, D., Karp-Boss, L., Roux, S., Sunagawa, S., Wincker, P., Sullivan, M.B., 2019. Marine DNA viral macro- and microdiversity from pole to pole. Cell 177, 1109, 1123.e1114.

Harrison, R.D., 2005. Figs and the diversity of tropical rainforests. Bioscience 55, 1053–1064.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Hou, X., He, Y., Fang, P., Mei, S.-Q., Xu, Z., Wu, W.-C., Tian, J.-H., Zhang, S., Zeng, Z.-Y., Gou, Q.-Y., Xin, G.-Y., Le, S.-J., Xia, Y.-Y., Zhou, Y.-L., Hui, F.-M., Pan, Y.-F., Eden, J.-S., Yang, Z.-H., Han, C., Shu, Y.-L., Guo, D., Li, J., Holmes, E.C., Li, Z.-R., Shi, M., 2023. Artificial intelligence redefines RNA virus discovery. bioRxiv 2023.04.18.537342.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386.

Jansen van Vuren, P., Wiley, M.R., Palacios, G., Storm, N., Markotter, W., Birkhead, M., Kemp, A., Paweska, J.T., 2017. Isolation of a novel orthobunyavirus from bat flies (*Eucampsipoda africana*). J. Gen. Virol. 98, 935–945.

Jiang, Z.-F., Huang, D.-W., Zhu, C.-D., Zhen, W.-Q., 2006. New insights into the phylogeny of fig pollinators using Bayesian analyses. Mol. Phylogenet. Evol. 38, 306–315.

Käfer, S., Paraskevopoulou, S., Zirkel, F., Wieseke, N., Donath, A., Petersen, M., Jones, T. C., Liu, S., Zhou, X., Middendorf, M., Junglen, S., Misof, B., Drosten, C., 2019. Re-assessing the diversity of negative strand RNA viruses in insects. PLoS Pathog. 15, e1008224.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Kamani, J., González-Miguel, J., Msheliza, E.G., Goldberg, T.L., 2022. Straw-colored fruit bats (*Eidolon helvum*) and their bat flies (*Cyclopodia greefi*) in Nigeria host viruses with multifarious modes of transmission. Vector Borne Zoonotic Dis. 22, 545–552.

King, A.M., Lefkowitz, E., Adams, M.J., Carstens, E.B., 2011. Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, Amsterdam, Netherlands.

Koonin, E.V., Dolja, V.V., Krupovic, M., Varsani, A., Wolf, Y.I., Yutin, N., Zerbini, F.M., Kuhn, J.H., 2020. Global organization and proposed megataxonomy of the virus world. Microbiol. Mol. Biol. Rev. 84, e00061-19.

Kuang, G., Xu, Z., Wang, J., Gao, Z., Yang, W., Wu, W., Liang, G., Shi, M., Feng, Y., 2023. Nelson Bay reovirus isolated from bats and blood-sucking arthropods collected in Yunnan Province, China. Microbiol Spectr, e0512222.

Kuhn, J.H., Koonin, E.V., 2023. Viriforms—a new category of classifiable virus-derived genetic elements. Biomolecules 13, 289.

Kunz, T.H., Braun de Torrez, E., Bauer, D., Lobova, T., Fleming, T.H., 2011. Ecosystem services provided by bats. Ann. N. Y. Acad. Sci. 1223, 1–38.

Kunz, T.H., Whitaker Jr., J.O., Wadanoli, M.D., 1995. Dietary energetics of the insectivorous Mexican free-tailed bat (*Tadarida brasiliensis*) during pregnancy and lactation. Oecologia 101, 407–415.

Lee, B.D., Neri, U., Roux, S., Wolf, Y.I., Camargo, A.P., Krupovic, M., Simmonds, P., Kyrpides, N., Gophna, U., Dolja, V.V., Koonin, E.V., 2023. Mining metatranscriptomes reveals a vast world of viroid-like circular RNAs. Cell 186, 646–661.e4.

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, e05378.

Manso, C.F., Bibby, D.F., Mbisa, J.L., 2017. Efficient and unbiased metagenomic recovery of RNA virus genomes from human plasma samples. Sci. Rep. 7, 4173.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534.

Neri, U., Wolf, Y.I., Roux, S., Camargo, A.P., Lee, B., Kazlauskas, D., Chen, I.M., Ivanova, N., Zeigler Allen, L., Paez-Espino, D., Bryant, D.A., Bhaya, D., RNA Virus Discovery Consortium, Krupovic, M., Dolja, V.V., Kyrpides, N.C., Koonin, E.V., Gophna, U., 2022. Expansion of the global RNA virome reveals diverse clades of bacteriophages. Cell 185, 4023–4037.e18.

Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 27, 824–834.

Olendraite, I., Brown, K., Firth, A.E., 2023. Identification of RNA virus-derived RdRp sequences in publicly available transcriptomic data sets. Mol. Biol. Evol. 40, msad060.

Ortiz-Baez, A.S., Jaenson, T.G.T., Holmes, E.C., Pettersson, J.H.-O., Wilhelmsson, P., 2023. Substantial viral and bacterial diversity at the bat-tick interface. Microb. Genom. 9, mgen000942.

Paraskevopoulou, S., Käfer, S., Zirkel, F., Donath, A., Petersen, M., Liu, S., Zhou, X., Drosten, C., Misof, B., Junglen, S., 2021. Viromics of extant insect orders unveil the evolution of the flavi-like superfamily. Virus Evol 7 veab030.

Parry, R., Asgari, S., 2018. Aedes anphevirus: an insect-specific virus distributed worldwide in *Aedes aegypti* mosquitoes that has complex interplays with *Wolbachia* and dengue virus infection in cells. J. Virol. 92, e00224-18.

Pascall, D.J., Tinsley, M.C., Clark, B.L., Obbard, D.J., Wilfert, L., 2021. Virus prevalence and genetic diversity across a wild bumblebee community. Front. Microbiol. 12, 650747.

Paskey, A.C., Ng, J.H.J., Rice, G.K., Chia, W.N., Philipson, C.W., Foo, R.J.H., Cer, R.Z., Long, K.A., Lueder, M.R., Frey, K.G., Hamilton, T., Mendenhall, I.H., Smith, G.J., Wang, L.-F., Bishop-Lilly, K.A., 2020. The temporal RNA virome patterns of a lesser dawn bat (*Eonycteris spelaea*) colony revealed by deep sequencing. Virus Evol 6, veaa017.

Peng, C., Zhang, D., Li, C., Li, Y., Zhang, H., Li, N., Xiao, P., 2022. *Rhinolophus sinicus* virome revealed multiple novel mosquito-borne zoonotic viruses. Front. Cell. Infect. Microbiol. 12, 960507.

Pettersson, J.H.-O., Shi, M., Bohlin, J., Eldholm, V., Brynildsrud, O.B., Paulsen, K.M., Andreassen, Å., Holmes, E.C., 2017. Characterizing the virome of *Ixodes ricinus* ticks from northern Europe. Sci. Rep. 7, 10870.

R Core Team, 2018. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rambaut, A., 2009. FigTree. http://tree.bio.ed.ac.uk/software/figtree/.

Ramírez-Martínez, M.M., Bennett, A.J., Dunn, C.D., Yuill, T.M., Goldberg, T.L., 2021. Bat flies of the family *Streblidae* (Diptera: hippoboscoidea) host relatives of medically and agriculturally important "bat-associated" viruses. Viruses 13, 860.
Sakuna, K., Elliman, J., Owens, L., 2017. Discovery of a novel *Picornavirales*, Chequa iflavirus, from stressed redclaw crayfish (*Cherax quadricarinatus*) from farms in northern Queensland, Australia. Virus Res. 238, 148–155.
Schoonvaere, K., De Smet, L., Smagghe, G., Vierstraete, A., Braeckman, B.P., de Graaf, D. C., 2016. Unbiased RNA shotgun metagenomics in social and solitary wild bees detects associations with eukaryote parasites and new viruses. PLoS One 11, e0168456.
Segura-Trujillo, C.A., Lidicker, W.Z., Álvarez-Castañeda, S.T., 2016. New perspectives on trophic guilds of arthropodivorous bats in North and Central America. J. Mammal. 97, 644–654.
Shi, C., Beller, L., Deboutte, W., Yinda, K.C., Delang, L., Vega-Rúa, 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., Chen, X., Tian, J.-H., Chen, L.-J., Li, K., Wang, W., Eden, J.-S., Shen, J.-J., Liu, L., Holmes, E.C., Zhang, Y.-Z., 2018. The evolutionary history of vertebrate RNA viruses. Nature 556, 197–202.
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.
Tisza, M.J., Pastrana, D.V., Welch, N.L., Stewart, B., Peretti, A., Starrett, G.J., Pang, Y.-Y. S., Krishnamurthy, S.R., Pesavento, P.A., McDermott, D.H., Murphy, P.M., Whited, J. L., Miller, B., Brenchley, J., Rosshart, S.P., Rehermann, B., Doorbar, J., Ta'ala, B.A., Pletnikova, O., Troncoso, J.C., Resnick, S.M., Bolduc, B., Sullivan, M.B., Varsani, A., Segall, A.M., Buck, C.B., 2020. Discovery of several thousand highly diverse circular DNA viruses. Elife 9, e51971.
Toohey-Kurth, K., Sibley, S.D., Goldberg, T.L., 2017. Metagenomic assessment of adventitious viruses in commercial bovine sera. Biologicals 47, 64–68.
Vainio, E.J., Chiba, S., Ghabrial, S.A., Maiss, E., Roossinck, M., Sabanadzovic, S., Suzuki, N., Xie, J., Nibert, M., ICTV Report Consortium, 2018. ICTV virus taxonomy profile: *Partitiviridae*. J. Gen. Virol. 99, 17–18.
Valles, S.M., Chen, Y., Firth, A.E., Guérin, D.M.A., Hashimoto, Y., Herrero, S., de Miranda, J.R., Ryabov, E., ICTV Report Consortium, 2017. ICTV virus taxonomy profile: *Iflaviridae*. J. Gen. Virol. 98, 527–528.
Vasilakis, N., Tesh, R.B., 2015. Insect-specific viruses and their potential impact on arbovirus transmission. Curr Opin Virol 15, 69–74.
Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theor. Biosci. 131, 281–285.
Webster, C.L., Longdon, B., Lewis, S.H., Obbard, D.J., 2016. Twenty-five new viruses associated with the Drosophilidae (Diptera). Evol Bioinform Online 12, 13–25.
Weiblen, G.D., 2002. How to be a fig wasp. Annu. Rev. Entomol. 47, 299–330.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis, second ed. Springer, New York, New York, USA.
Xu, Z., Feng, Y., Chen, X., Shi, M., Fu, S., Yang, W., Liu, W.J., Gao, G.F., Liang, G., 2022. Virome of bat-infesting arthropods: highly divergent viruses in different vectors. J. Virol. 96, e0146421.
Ye, Z.-X., Li, Y.-H., Chen, J.-P., Zhang, C.-X., Li, J.-M., 2021. Complete genome analysis of a novel iflavirus from a leaf beetle, *Aulacophora lewisii*. Arch. Virol. 166, 309–312.
Yinda, C.K., Zell, R., Deboutte, W., Zeller, M., Conceição-Neto, N., Heylen, E., Maes, P., Knowles, N.J., Ghogomu, S.M., Van Ranst, M., Matthijnssens, J., 2017. Highly diverse population of *Picornaviridae* and other members of the *Picornavirales*, in Cameroonian fruit bats. BMC Genom. 18, 249.
Zayed, A.A., Wainaina, J.M., Dominguez-Huerta, G., Pelletier, E., Guo, J., Mohssen, M., Tian, F., Pratama, A.A., Bolduc, B., Zablocki, O., Cronin, D., Solden, L., Delage, E., Alberti, A., Aury, J.M., Carradec, Q., da Silva, C., Labadie, K., Poulain, J., Ruscheweyh, H.J., Salazar, G., Shatoff, E., Tara Oceans Coordinators, Bundschuh, R., Fredrick, K., Kubatko, L.S., Chaffron, S., Culley, A.I., Sunagawa, S., Kuhn, J.H., Wincker, P., Sullivan, M.B., Acinas, S.G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., de Vargas, C., Gorsky, G., Guidi, L., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Kandels, S., Karp-Boss, L., Karsenti, E., Not, F., Ogata, H., Poulton, N., Pesant, S., Sardet, C., Speich, S., Stemmann, L., Sullivan, M.B., Sungawa, S., Wincker, P., 2022. Cryptic and abundant marine viruses at the evolutionary origins of earth's RNA virome. Science 376, 156–162.
Zhao, G., Wu, G., Lim, E.S., Droit, L., Krishnamurthy, S., Barouch, D.H., Virgin, H.W., Wang, D., 2017. VirusSeeker, a computational pipeline for virus discovery and virome composition analysis. Virology 503, 21–30.
Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., Alva, V., 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol. 430, 2237–2243.