Received: 13 June 2024 | Revised: 25 July 2024 | Accepted: 15 August 2024

DOI: 10.1111/mec.17519

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Viral diversity and co-evolutionary dynamics across the ant phylogeny

Peter J. Flynn[1,2] | Corrie S. Moreau[3,4]

[1]Committee on Evolutionary Biology, University of Chicago, Chicago, Illinois, USA

[2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

[3]Department of Entomology, Cornell University, Ithaca, New York, USA

[4]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York, USA

Correspondence
Peter J. Flynn, Department of Organismic and Evolutionary Biology, Harvard University, 22 Divinity Avenue, Cambridge, MA 02138, USA.
Email: pflynn@fas.harvard.edu

Funding information
Division of Biological Infrastructure, Grant/Award Number: 2210800; Division of Environmental Biology, Grant/Award Number: 1900357; National Science Foundation Graduate Research Fellowship Program

Handling Editor: Kayla King

Abstract

Knowledge of viral biodiversity within insects, particularly within ants, is extremely limited with only a few environmental viruses from invasive ant species identified to date. This study documents and explores the viral communities in ants. We comprehensively profile the metagenomes of a phylogenetically broad group of 35 ant species with varied ecological traits and report the discovery of 3710 novel and unique ant-associated viral genomes. These previously unknown viruses discovered within this study constitute over 95% of all currently described ant viruses, significantly increasing our knowledge of the ant virosphere. The identified RNA and DNA viruses fill gaps in insect-associated viral phylogenies and uncover evolutionary histories characterized by both frequent host switching and co-divergence. Many ants also host diverse bacterial communities, and we discovered that approximately one-third of these new ant-associated viruses are bacteriophages. Two ecological categories, bacterial abundance in the host and habitat degradation are both correlated with ant viral diversity and help to structure viral communities within ants. These data demonstrate that the ant virosphere is remarkably diverse phylogenetically and genomically and provide a substantial foundation for studies in virus ecology and evolution within eukaryotes. We highlight the importance of studying insect-associated viruses in natural ecosystems in order to more thoroughly and effectively understand host-microbe evolutionary dynamics.

KEYWORDS
biodiversity, co-evolution, ecology, Formicidae, host-microbes, viruses

1 | INTRODUCTION

Ants (Hymenoptera: Formicidae) are one of the most diverse and abundant groups of organisms on Earth (Mora et al., 2011), and are a highly successful clade in terms of biomass and species diversity. Their associated microbes play an important functional role for species of varying dietary types. Microbial symbionts have been theorized to largely enable the ecological dominance of ants throughout nutrient-limited rainforest canopies (Cook & Davidson, 2006; Davidson et al., 2003). Other molecular surveys have illustrated that symbiotic bacteria are correlated with the evolution of herbivory across ant species (Russell et al., 2009), while both experimentation and genomics have demonstrated that herbivorous ants benefit from specialized, nitrogen recycling gut bacteria (Bisch et al., 2018; Duplais

et al., 2021; Feldhaar et al., 2007; Gil et al., 2003; Hu et al., 2018). To date, however, little is known about ant viruses, with only a small number of ant-associated viruses ever identified (Baty et al., 2020).

Ants have been known to be infected by viruses since the mid-1900s (Steiger et al., 1969). However, very few of the over 14,000 species of ants have even been sampled for potential viral infection. Only 40 unique viruses have been found within 38 ant species to date, mostly from invasive ant species, such as *Solenopsis invicta* (red imported fire ant) and *Linepithema humile* (Argentine ant) with the most well studied being the Solenopsis Invicta Viruses (SINV-1-11; Olendraite et al., 2017; Sébastien et al., 2015; Valles et al., 2004, 2014, 2018). This bias derives from more intensive sampling of invasive ants as their viruses are potential targets for biocontrol agents and because they may act as potential reservoirs for detrimental honey bee viruses (Lester et al., 2019). The most common viruses found across ant species are associated with and were first discovered within honey bees (Baty et al., 2020). There appears to be frequent transmission between honey bee hives and ant colonies due to ant raids on bee colonies. Though these viruses are often thought of as 'honey bee viruses', they have been observed to infect many other hymenopteran hosts and may have evolved in non-honey bee hymenopterans and jumped hosts over time (Loope et al., 2019).

While it is now possible to discover eukaryotic host-associated viral communities using metagenomic techniques (Lim et al., 2015; Paez-Espino et al., 2016; Roux et al., 2017; Shi et al., 2016; Vibin et al., 2018), we still have an incomplete understanding of the factors that structure viral communities within eukaryotes. Metagenomic studies have demonstrated that features of host biology and ecology can significantly impact virus association and therefore could also be major drivers of viral emergence (Geoghegan et al., 2021; Wille, 2020). Here we describe a large-scale metagenomic survey of diverse ant hosts to reveal the unprecedented diversity of DNA and RNA viruses within ants. This study is the first extensive characterization of ant-associated viruses. We investigated the composition and diversity of viral communities of a phylogenetically broad group of ant species across a pristine Amazonian rainforest and fragmented urban habitats in South America. We tested whether certain viral communities tend to associate with specific ant species depending on the differences in those ant species' ecology, diet, bacterial abundance, phylogeny and habitat type. This species and habitat specific data enables us to better explain the origin and evolutionary history of ant-associated viruses as well as begin to uncover their fundamental role in the patterns and processes of ant evolution.

2 | MATERIALS AND METHODS

During a 2-week period in March 2018, ants were sampled from two locations, in a pristine rainforest within Nouragues Reserve, French Guiana and within the fragmented urban habitat around Cayenne, French Guiana (Figure S1). We collected ant species found in French Guiana each with a paired sister species when applicable to be able to control for close phylogenetic relationships. Only adult worker ants were collected from each colony to control for caste-specific variability. Worker subtypes (forager and nurse) were not assessed for any of the ants sampled. As many worker ants as was feasible from the same colony were collected (ranging from 4 to 75). Species were identified on site, using external morphological characteristics. The samples were brought alive to the Pasteur Institute in Cayenne. The ants were fed the same diet (sterile sugar water) for a week before freezing so that viruses found in the gut would not be due to diet (i.e. fungal or bird faeces-associated viruses). A voucher specimen for each colony is deposited in the Cornell University Insect Collection, in Ithaca, New York, USA.

Each sample was pooled for a total of 44 samples and an additional control sample. Each ant from the same colony was placed alive into a 50 mL Falcon™ tube and stored at −80°C. Each ant was individually washed with 99.95% ethanol, 10% bleach, and washed three times with sterile nuclease-free water to clean the cuticle, to avoid external viral contamination. The samples were flash frozen with liquid nitrogen and crushed manually. Samples were homogenized with 10 mL DMEM and cleared of debris by centrifugation (5 min, 10,000 *g*, 4°C). Prokaryotic and eukaryotic cell-sized particles were removed from the supernatant through successive filtrations (0.8, 0.45, 0.22 μm), using cellulose–acetate membrane filters (Nalgene). Filtrates were cleared of less dense components through a 1-h ultracentrifugation (100,000 *g*, 4°C), then resuspended and cleared of persistent high-density particles with centrifugation (15 min, 10,000 *g*, 4°C). Viral particles were pelleted with an additional 1 h ultracentrifugation (Salmier et al., 2017).

Because a diversity of contaminants, including viruses can be found in laboratory reagents and commercially available extraction/amplification kits, a negative control sample containing 0.5 mL HBSS-1X alone was included. This negative control was processed alongside the other samples from the ant washing step onward to control for unexpected contamination in reagents which could potentially influence metagenomic analyses of low biomass samples (Rosario, Fierer, et al., 2018).

Resuspended viral pellets were treated with a mixture of Turbo DNase (Ambion), Benzonase (Novagen) and RNase-One (Promega) to digest nonenveloped nucleic acids (Allander et al., 2001). Viral DNA and RNA was extracted and purified using the QIAmp Viral-RNA mini kit (Qiagen) without carrier RNA. For each sample, DNA virus-only and RNA virus-only libraries were constructed using a whole transcriptome or whole genome amplification method (Berthet et al., 2008). For the RNA virus-only amplification, an aliquot of half the nucleic acid collected was treated with Turbo DNase to remove viral DNA. Persistent rRNA was depleted with GeneRead rRNA-depletion kit (Qiagen). Viral RNA amplification was performed using the QuantiTect Whole Transcriptome Kit (Qiagen). cDNA was synthesized using SuperScript1-III Reverse Transcriptase (Invitrogen) and random hexamers (Roche). For DNA virus-only amplification, an aliquot of half the extracted viral nucleic acids was treated with RNase-One to remove viral RNA. Then viral DNA amplification was performed using the QuantiTect Whole Genome Kit (Qiagen). To ensure homogeneity and blunt ends, 10 U of Klenow polymerase

(Roche) was added to the amplified nucleic acids with 8 μL of random hexamers (Roche), incubated 1 h at 37°C, then 10 min at 75°C. Samples were quantified using dsDNA Qubit assay kits (Qiagen).

Each sample was pooled for a total of 44 samples and an additional control sample. All sequencing was performed at Centre INRA in Toulouse, France. The genomic sequencing libraries was generated using the TruSeq DNA PCR-Free Sample Preparation Kit (Illumina). Sequencing was carried out on a HiSeq3000 sequencing platform (Illumina) using paired-end (2 × 150 bp) reads across three runs.

The raw sequencing reads from each sample library were concatenated from all three runs. Raw reads were trimmed for quality and adaptors removed using Trimmomatic v0.36 with default parameters (Bolger et al., 2014). The quality of the trimmed reads was verified using FastQC (Simons, 2010). Quality-filtered reads were mapped to the most closely related ant species genome using Bowtie2 to remove ant DNA contamination (Langmead & Salzberg, 2012). For each sample library, the reads which did not map to an ant genome were de novo assembled using SPAdes-v3.14.0 (Bankevich et al., 2012). For the SPAdes assembly, the single-cell mode was used with K-mer sizes set at 21, 33, 55, 77, 99 and 127. Contigs larger than 500 bp were retained. Assembled contigs within each sample were dereplicated based on a 95% identity cutoff using CD-HIT (Fu et al., 2012). Contaminant contigs were removed from each sample sequence library by comparing BLASTn (*e*-value <0.00001) against a database containing assembled sequences from the negative control library (Rosario, Fierer, et al., 2018). Additionally, CheckV-0.8.1 was used to remove contigs suspected of host-derived sequence proviral contamination (Nayfach et al., 2021).

To identify viral sequences, contigs from each library were compared (BLASTx, *e*-value <0.001) against a viral protein database containing sequences from NCBI RefSeq database using DIAMOND (Buchfink et al., 2014). To further determine the prokaryotic viruses in these samples, contigs were analysed using VirSorter-2.2.1 (Guo et al., 2021). Contig sequences with significant matches to the viral RefSeq database and VirSorter were compared against the GenBank nonredundant database (BLASTx, *e*-value<0.001, downloaded: 11/02/22) to remove sequences that had higher identity with non-viral sequences. To approximate viral taxonomy, the most closely related viral taxa were defined using best-hit filtering.

The putative viral sequences were then sorted by taxonomy of top viral hit from the non-redundant protein database. EMBOSSv6.6.0 predicted viral open reading frames for these putative viral sequences (Rice et al., 2000). The putative viral amino acid (AA) sequences were compared (BLASTp, *e*-value <0.001) to the nonredundant protein database to determine viral proteins which definitively matched viral hallmark proteins within their taxonomic group. Any contig which contained a viral hallmark protein was henceforth termed 'viral genome fragments' (VGF) or 'phage genome fragment' (PGF) and included in subsequent analyses.

These putative viral protein sequences were aligned with all related viruses from the same viral clades (downloaded from NCBI) using MAFFT v7.309 employing the E-INS-i algorithm (Katoh et al., 2002). ProtTest-v3.4.2 determined the best-fit AA substitution model for each alignment (Darriba et al., 2011). Subsequently, maximum likelihood viral phylogenies were inferred using the nucleotide alignments with RAxML-v8.2.11 (Stamatakis, 2014). These phylogenies were used to infer the evolutionary relationships between these putative viral sequences and their closest viral relatives (Shi et al., 2016). Any newly discovered partial viral genome that clustered within the specific viral phylogeny was labelled a putative ant-associated VGF/PGF.

Completeness of these VGFs were assessed using CheckV-0.8.1 and manually evaluated based on the most closely related viral lineage and viral proteins compared with ORFs within the viral contig, with annotation performed in Geneious Prime-2020.0.4 (https://www.geneious.com), using all available genomes within the viral clade as reference (Nayfach et al., 2021). When VGFs of these ant samples had >95% AA similarity to their top viral protein hit they were considered the same viral species. If the entire viral genome had >80% nucleotide pairwise similarity to their top viral genome hit, they were considered an isolate of that viral genome. Assembled contigs *between* each sample were considered duplicates of each other if they clustered together based on a 95% sequence-wide identity cutoff using CD-HIT (Fu et al., 2012). Cressdnaviricota (CRESS) genomes were inspected for the presence of three RCR and superfamily-3 (SF3) helicase motifs (Rosario, Dayaram, et al., 2012; Rosario, Duffy, & Breitbart, 2012; Zuker, 2003). We used ggplot2 for all figure visualizations and ITOL for phylogenetic visualizations (Letunic & Bork, 2007; Wickham, 2016).

Co-phylogeny analyses were performed using three methods: Procrustes approach to co-phylogeny (PACo), Jane4 software and Bayesian tip-association significance testing software (BaTS) (Conow et al., 2010; Hutchinson et al., 2017; Parker et al., 2008). PACo assesses the evolutionary dependency of two groups of interacting species using their phylogenetic history. The PACo method was employed using *Paco* in R to test for significant congruence between host–parasite phylogenies (Balbuena et al., 2013; Hutchinson et al., 2017; R Core Team, 2017). For PACo analyses, the null model selected was r0, which assumes that virus phylogeny tracks host phylogeny. Levels of cophylogenetic signal were evaluated as the median global sum-of-squared residuals (m^2_{xy}) and mean significance averaged over 100,000 posterior trees. Jane is an event-based co-phylogenetic reconstruction approach, which evaluates the extent of virus–host co-divergence within each viral clade (Conow et al., 2010). Only ant-associated viruses identified in this study were included in these viral phylogenies. For the host phylogeny, we used the ant phylogeny from Nelsen et al. (2018), including only the 35 ant species sampled (Nelsen et al., 2018). The cost-scheme was set to default and generations and population-size to 500. The co-divergence significance was derived by comparing the estimated costs of null distributions from 1000 randomizations of host tip mapping. We visualized these associations between the viral and ant phylogeny using *phytools* to create a tanglegram (Revell, 2012). Non-random trait association of tips across the viral phylogeny was tested using BaTS (Parker et al., 2008; Shi et al., 2018). BaTS considered ant host phylogenetic structure at species, genus and subfamily, then

estimated an association index (AI) to identify the strength of the association between viral-ant host phylogenies and compared to a null distribution (over 10,000 tree-tip randomizations) to infer an AI-Ratio (observed AI/null AI). We chose to include the bacteriophage lineages (Microviridae and Caudoviricetes) since previous studies have illustrated co-diversification patterns of bacteriophage to their animal hosts (Wu et al., 2024).

BaTS was also used to test for the association between ecological traits of the ant host species and every inferred viral phylogeny. The ecological traits tested were diet (herbivorous, carnivorous and omnivorous), habitat (urban and rainforest), nest type (ground, arboreal or leaf litter/rotten log) and bacterial abundance within ant colony measured through qPCR (low, medium and high). All of these ecological traits were described for each of these ant samples in Chanson et al. (2023). AI was estimated to identify the strength of the association between the viral phylogeny and particular ecological trait of the ant host. Because all of these ant ecological traits were categorical variables, δ statistic was calculated to evaluate the effects of phylogeny on the evolution of a categorical trait across species (Borges et al., 2019; Wang et al., 2021). The estimated δ for the given trait is compared to the null distribution for random δ (over 10,000 iterations) to assess significant phylogenetic signal for the trait. The δ statistic decreases when the trait evolved independently across the phylogeny and the δ value was calculated in *ape* within R (Paradis & Schliep, 2019).

3 | RESULTS

3.1 | Ant-associated viral identification

A total of 35 ant species from 44 ant colonies and one control were sequenced for DNA and RNA viruses (Table S1). Multiple worker ants were collected from the same colony, ranging from four to 75 ants per colony depending on worker size (Table S1). Additionally, there were nine identical ant species paired between the urban and rainforest habitats (Table S1 and Figure S1). The total raw sequencing reads were 769,125,840 reads, with an average of 17,091,685 reads per ant colony sample. There was variability in the number of raw sequencing reads with the most being *Paraponera clavata* (CSM3708) with 22,544,316 reads and the least, *Odontomachus haematodus* (CSM3670) with 6,178,730 reads. Though we recorded read abundance for every ant sample, it was not analysed more thoroughly due the multiple displacement amplification (MDA) methodology, which can lead to biases towards circular ssDNA viruses and skew read abundance estimates (Sullivan et al., 2016).

A total of 3841 ant-associated viral genomes (both partial and complete) were identified in this study, with 3710 of these viruses considered both unique and novel. Based on sequence similarity, only 52 viruses within the ant samples were considered the same viral species as their top viral protein hit, and therefore not 'novel'. Additionally, there were 72 clusters of both partial and complete viral genomes which had >95% sequence identity to each other and were therefore considered identical and not 'unique' (25 Microvirus, 22 CRESS virus, 17 Parvovirus, 6 Picorna-Calici, 1 Caudoviricetes and 1 Crucivirus clusters- constituting 79 'not unique' viruses). These identical clusters consisted of viruses primarily from ants within different subfamilies, though there were several instances of identical viruses found within subfamily, genus, and species level (Data S1). Of the 3841 viruses identified, a total of 659 viral genomes were assessed as complete (17.20% of total; Table 1).

The most common VGFs identified were most closely related to the Microviridae viral clade (1339), Picorna-Calici clade (1135), CRESS (637), Parvoviridae (249) and Caudoviricetes (135) (Figure 1). Though the majority of these VGFs fell within these five viral clades, there was a wide overall diversity with 13 viral RNA clades and nine viral DNA clades represented. Additionally, there was variability between which ant genera hosted the largest number of VGFs (Figure 2). Within our study, *Dolichoderus* (882 viruses), *Ectatomma* (784), *Crematogaster* (573), *Daceton* (355) and *Odontomachus* (305) all hosted the largest number of viruses, whereas several genera including *Atta* (25), *Pseudomyrmex* (19) and *Anochetus* (3) hosted far few viruses (Figures S2 and 3). This same pattern held true when identifying viruses by ant species (Figure 3). Colony level viral patterns could be impacted by the fact we did not classify by worker subtype (i.e. foraging workers could harbour more viruses than nursing workers). The viral genome type appeared to be non-randomly distributed, with 58.97% of the VGFs derived from ssDNA viruses and 34.55% from ssRNA+ viruses, whereas ssRNA-, dsRNA, and dsDNA all comprised the remainder (Figure S3). This non-random distribution could in part be due to the MDA bias towards sequencing small, circular ssDNA viruses (Sullivan et al., 2016). Before this study there were a total of 170 viruses previously described in ant species, with 40 unique viruses described (Figure 1). Of these 170 viruses, the majority were described from the ant genus, *Solenopsis* (29.41%) and other invasive ant species.

Though we identified viruses within both Bidnaviridae and Nucleocytoviricota, however, since none of the viruses in our dataset definitively matched viral hallmark proteins, we did not infer phylogenies for either of these clades. Additionally, if the specific viral genome did not include the hallmark protein of its viral clade it was not included in a phylogeny. Therefore, a totalled of 120 viruses within our dataset were not included in the inferred viral phylogenies. All phylogenies from every inferred viral clade discussed below can be found at https://doi.org/10.5281/zenodo.11643759.

3.2 | Novel positive-sense single-stranded RNA (+ssRNA) viruses

3.2.1 | Astro

Two VGF both from *Dolichoderus attelaboides* were most closely related to the Astrovirus clade (Figure 2a). Dolichoderus attelaboides Astro VGF-1 fell within the passerine astrovirus clade, whereas Dolichoderus attelaboides Astro VGF-2 was most closely related to the Hainan gecko similignum astrovirus, within a group of vertebrate viruses.

TABLE 1 Table of viral genomes.

Viral lineage	Genome type	Total viral contigs	Complete genomes	Used	Hallmark protein	Substitution model
Anelloviridae	ssDNA	1	1	1	VP2	VT+I+G
Astro	ssRNA+	2	0	2	RdRp	LG+I+G
Bunya-Arena	ssRNA−	20	0	20	RdRp	VT+G+F
Caudoviricetes	dsDNA	135	1	58	TerL	Blosum62
Bidnaviridae	ssDNA	5	0	NA	NS1	NA
Cressdnaviricota (CRESS)	ssDNA	637	181	637	Rep	VT+G
Cruciviridae	ssDNA	34	15	34	Rep	Blosum62+G
Hepe-Virga	ssRNA+	8	0	8	RdRp	LG+G
Lefavirales	dsDNA	8	2	3	PIF1	WAG+I+G+F
Luteo-Sobemo	ssRNA+	70	0	70	RdRp	LG+I+G
Microviridae	ssDNA	1339	405	1339	MCP	PMB+G+F
Mono-Chu	ssRNA−	7	0	7	RdRp	LI+I+G
Narna-Levi	ssRNA+	34	1	34	RdRp	VT+G
Nucleocytoviricota	dsDNA	3	0	NA	MCP	NA
Partiti-Picobirna	dsRNA	35	0	35	RdRp	BLOSUM62+G
Parvoviridae	ssDNA	249	53	219	NS1	LG+G
Permutotetra	ssRNA+	33	0	33	RdRp	VT+G
Picorna-Calici	ssRNA+	1135	0	1135	RdRp	PMB+G
Reo	dsRNA	2	0	2	RdRp	LG+G
Tombus-Noda	ssRNA+	45	0	45	RdRp	LG+G
Toti-Chryso (including Mycovirus)	dsRNA	33	0	33	RdRp	Blosum62+G
Weivirus	dsRNA	6	0	6	RdRp	Blosum62+I+G
	Total	3841	659	3721		

Note: 'Total viral contigs' are the number of total viral genome fragments combined with the complete viral genomes discovered within this study and the complete genomes are the number of complete viral genomes discovered. 'Used' are the number of viruses used in the phylogenetic inference for this viral clade. The hallmark protein column is the viral protein identified in the literature which is most frequently used for phylogenetic inference. The 'substitution model' column represents the best-fit amino acid substitution model for each viral alignment. More detailed information on the specific viral contigs discovered within this study can be found in Data S1.

3.2.2 | Hepe-Virga

Eight ant-associated VGF fell within the Hepe-Virga RdRp phylogeny across three species (*Atta cephalotes*, *Camponotus femoratus*, *Nylanderia* sp. and *Dolichoderus attelaboides*; Figure 2g). Interestingly, five of these VGF (Atta cephalotes Hepe-Virga VGF-1 and -2, Dolichoderus attelaboides Hepe-Virga VGF-1, and Camponotus femoratus Hepe-Virga VGF-1) were demonstrated to be closely related to plant viruses in clades which serve as natural hosts for plants such as Virgaviridae, Tymoviridae and Betaflexviridae. These putative ant-associated viruses might be insect vectors for plant pathogenic transmission. Nylanderia sp. Hepe-Virga VGF-1 and Dolichoderus attelaboides Hepe-Virga VGF-2 fell within fungal-associated viruses within the Benyvirus clade. Nylanderia sp. Hepe-Virga VGF-2 clustered within a clade of Virgaviridae viruses which infect insects. Additionally, Nylanderia sp. Hepe-Virga VGF-3 fell next to Hepeviridae viruses associated with the faeces of vertebrates which prey on ants (*Vulpes vulpes* and *Gymnophithys rufigula*).

3.2.3 | Luteo-Sobemo

There were 70 ant-associated VGF within the Luteo-Sobemo clades across 15 ant species (Figure 2i). Plants serve as natural hosts for Luteo-Sobemo viruses, however many of these viruses use insects as transmission vectors. Most of the ant viruses discovered fell within the Sobemovirus clade (47 Luteo-Sobemo VGF), however two VGF clustered with Barnaviruses (Dolichoderus attelaboides Luteo-Sobemo VGF-1 and Camponotus femoratus Luteo-Sobemo VGF-1), six VGF were most closely related to Poleroviruses. Additionally, 15 VGF across seven ant species formed a sister clade to Poleroviruses, possibly representing ancient divergence of ant Poleroviruses.

3.2.4 | Tombus-Noda

Of the 45 Tombus-Noda associated ant VGF, most clustered with Tombusviridae (36 ant VGF) across 17 ant species (Figure 2r). Tombusviridae generally infect plants, however recently several

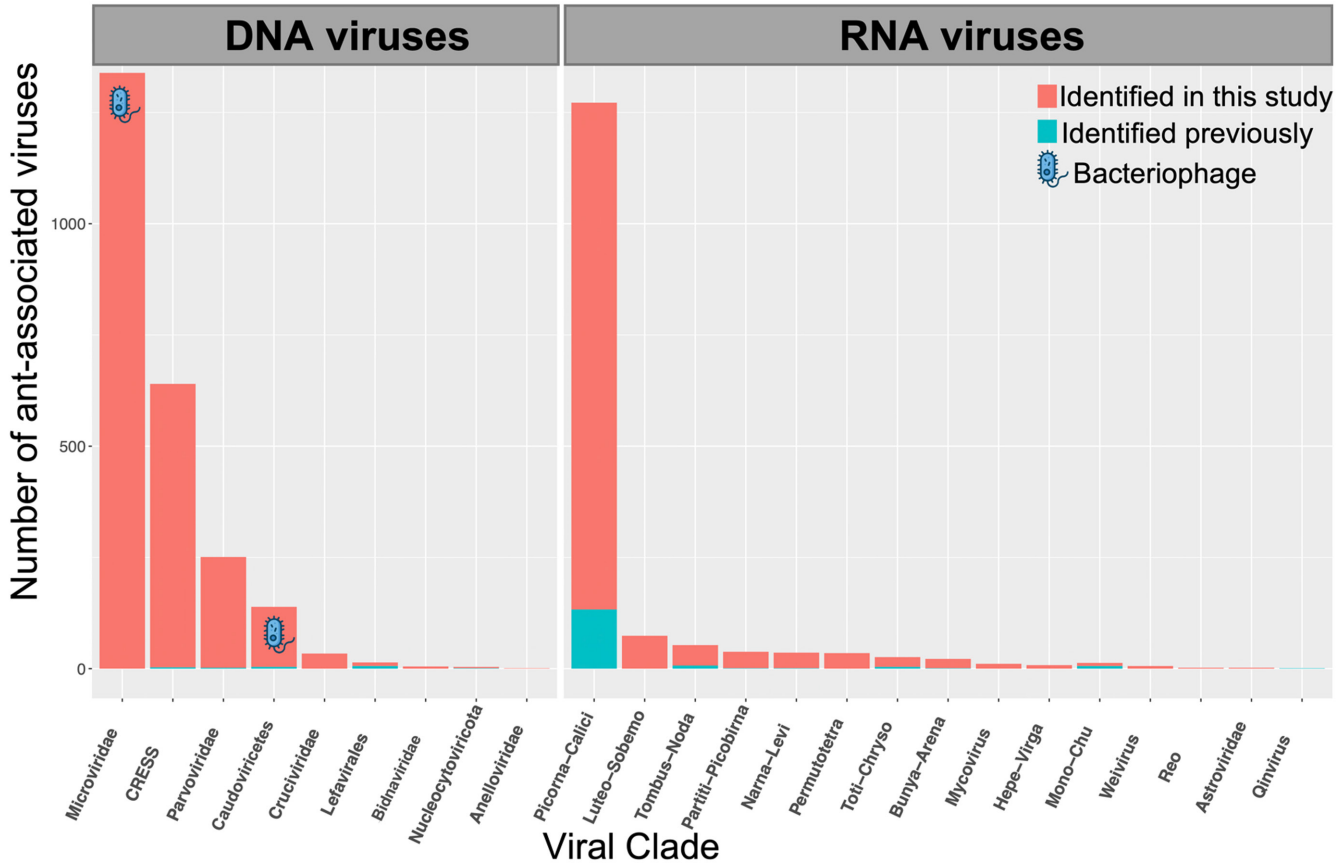


FIGURE 1 Identification of ant-associated viruses across divergent viral clades within widespread ant host genera. Total number of viruses within each viral clade. Red and blue represent current and previously discovered ant-associated viruses, respectively. Bacteria symbol represents viruses discovered which are known to infect bacteria.

Tombus-associated viruses have been demonstrated to infect insects. Nodaviridae generally infect vertebrates, yet the nine ant-associated VGF clustered with several insect-associated ant viruses within the Nodaviridae clade (such as Gungahlin Chrysomya nodalike virus, Sano virus and Carano virus).

3.2.5 | Permutotetra

There were 33 total Permutotetra-associated ant VGF (Figure 2o). Three of these viruses formed a distinct cluster outside of established Permutotetra viral lineages (Dolichoderus attelaboides Permutotetra VGF-1, Camponotus ant VGF-4 and Camponotus femoratus Permutotetra VGF-5). Insects are the natural hosts of Permutotetra viruses and the other 30 ant-associated Permutotetra VGF (across 11 ant species) clustered around the previously discovered insect-associated Permutotera.

3.2.6 | Narna-Levi

Within the Narna-Levi phylogeny there were 33 ant-associated Narna-Levi VGF and one complete Cephalotes atratus Narna-Levi virus 1, which is the only complete RNA virus in this dataset (Figure 2l). The Narna-Levi viruses within this dataset were mostly associated with *Cephalotes* and *Crematogaster* ant hosts (27/34). 15 *Cephalotes* and *Crematogaster* associated VGF (and Cephalotes atratus Narna-Levi virus 1) fell outside of classified Narna-Levi viruses. The rest of the ant-associated VGF were spread throughout the Narna-Levi phylogeny.

3.2.7 | Picorna-Calici

Picorna-Calici were by far most prevalent ant-associated RNA viruses with 1135 VGF across 30 species of ants (Figure 2p). 918 ant-associated VGF fell within unclassified Picornavirales and created an ant-specific clade without clustering with any other previously discovered viruses, possibly representing ancient divergence of Picorna-Calici ant-associated viruses. 74 additional ant-associated VGF were fell within unclassified Picornavirales viruses. 31 ant-associated Picorna-Calici VGF were most closely related to Caliviridae viruses. Dolichoderus attelaboides Picorna-Calici VGF-60 was found within Paavivirinae. Daceton armigerum Picorna-Calici VGF-208 fell within Kobuvirus. Dolichoderus attelaboides Picorna-Calici VGF-19 clustered with Sicinivirus. One ant-associated VGF fell within the lineage Kodimesavirinae - Crematogaster ant Picorna-Calici VGF-45. Within

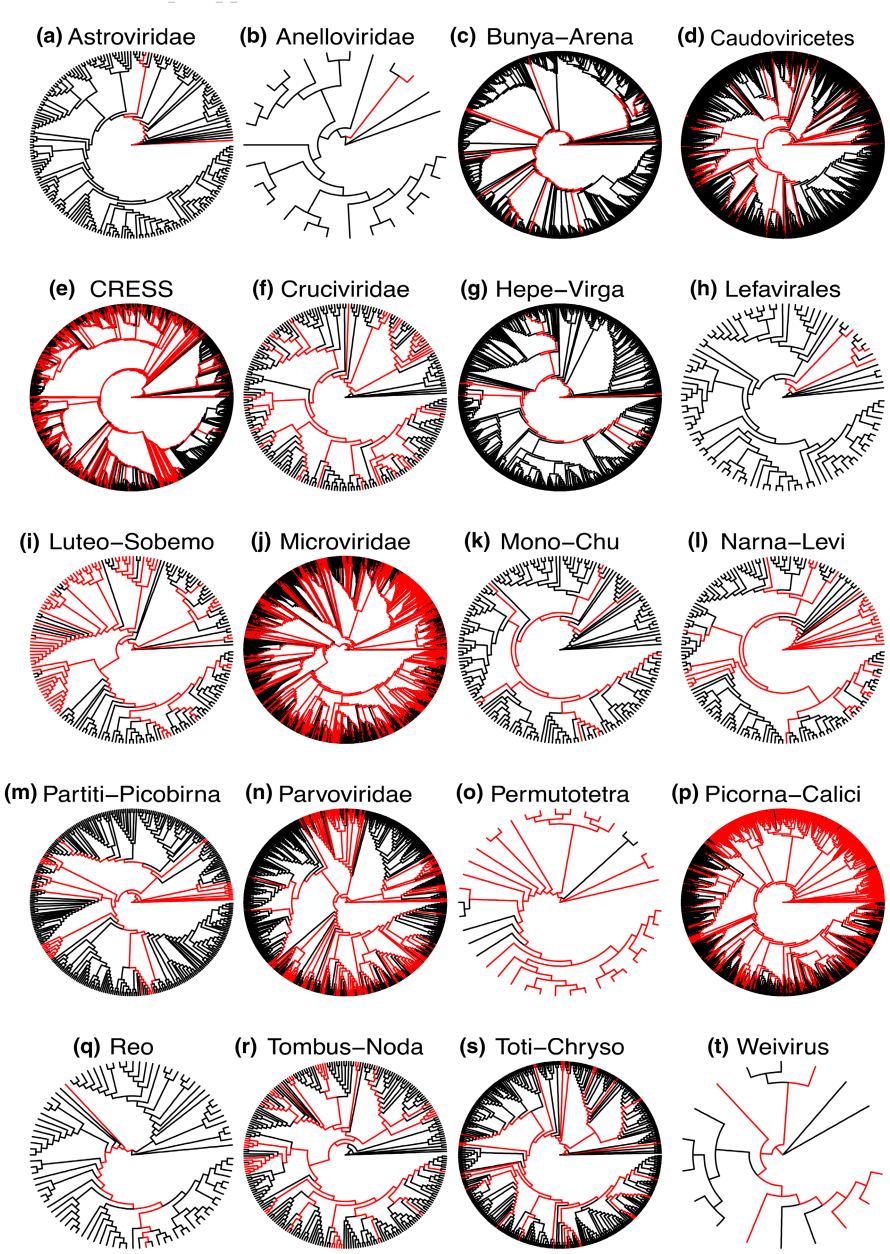


FIGURE 2 Twenty phylogenetic trees representing the major clades of virus discovered in this study. Cladograms a–t are midpoint rooted and within each tree, the viruses discovered in this study are shaded red, while every virus on NCBI found within that specific viral clade are shaded in black. Detailed phylogeny files with sequence names and bootstrap supports are available at Zenodo.

Caphthovirinae, there were two ant-associated VGF (Solenopsis geminata Picorna-Calici VGF-11 and Cephalotes ant Picorna-Calici VGF-23).

Seven ant-associated VGF fell within the Ensavirinae clade. Within Iflaviridae there were 39 closely related ant-associated VGF. 60 ant-associated VGF clustered with the Dicistroviridae family within Picorna-Calici.

3.3 | Novel negative-sense single-stranded RNA (-ssRNA) viruses

3.3.1 | Bunya-Arena

A total of 20 ant-associated VGF were identified as most closely related to the exogenous Bunya-Arena viral clade (Figure 2c). The VGF came from *Odontomachus haematodus* (13 from the same colony), and *Crematogaster* ants (six from two species), as well as one VGF from *Daceton armigerum*. Overall, these VGF were quite evenly distributed across the Bunya-Arena RdRp phylogeny, with a cluster of five VGF (Odontomachus haematodus Bunya-Arena VGF-3, -4, -6 and -12, and Crematogaster levior Bunya-Arena VGF-3) within the Phasmaviridae lineage, which are a clade of viruses which primarily infect insects.

3.3.2 | Mono-Chu

Seven VGF fell within the Mono-Chu lineage, which are known to commonly infect both vertebrates and invertebrates (Figure 2k). Cephalotes ant Mono-Chu VGF-2 and Daceton armigerum

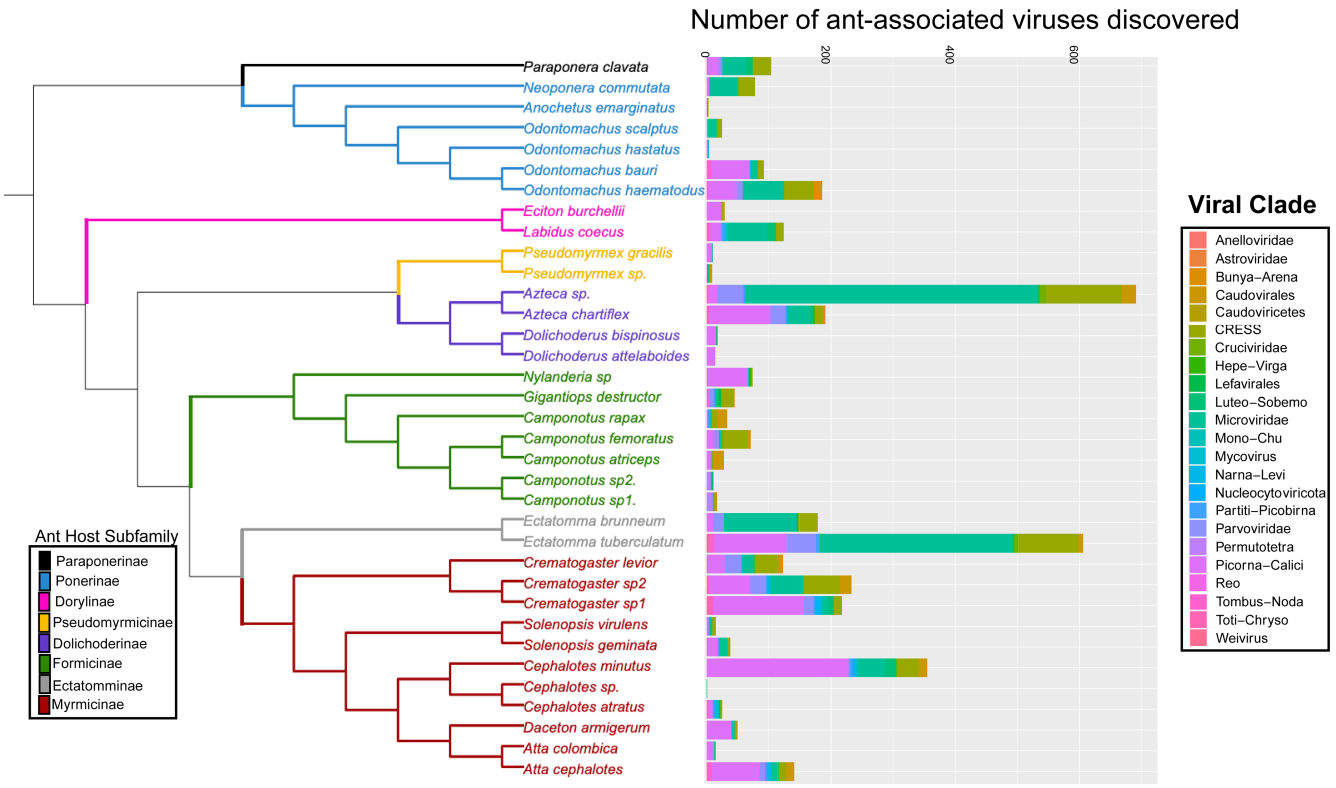


FIGURE 3 Phylogenetic relationships of ant hosts and viral genome fragments discovered. Phylogeny of the ant host species surveyed modified from (Nelsen et al., 2018), coloured by ant subfamily. Bar graph represents number of ant-associated viruses discovered per ant species, coloured by viral clade.

Mono-Chu VGF-2 were found to be most closely related to bee-associated viruses within Rhabdoviridae. Cephalotes ant VGF-1 fell within Mymonaviridae where fungus serve as the predominant host, though also recently been found in insects and plants. Daceton armigerum Mono-Chu VGF-1 and -3 clustered with Orinoviruses which frequently infect insects and Daceton armigerum Mono-Chu VGF-4 sister to Chuviridae which are associated with many species of insect. Cephalotes ant Mono-Chu VGF-3 was sister to a large clade of Mononegavirales viruses.

3.4 | Novel double-stranded RNA (dsRNA) viruses

3.4.1 | Toti-Chryso

Fungus serves as natural hosts of Toti-Chryso viruses, though recently there has been discovery of Toti-Chryso viruses infecting insect, plants, crustaceans, and mammals. There were 33 ant-associated Toti-Chryso VGF discovered within this study (Figure 2s). The vast majority of the Toti-Chryso VGF (31) were most closely related to Totiviridae viruses. Within the Totiviridae lineage, Crematogaster ant Toti-Chryso VGF-2 and -8 fell within the Victorivirus genus, which currently only has fungal hosts. Two VGF-Labidus coecus Toti-Chryso VGF-3 and Ectatomma tuberculatum Toti-Chryso VGF-2 fell within Chrysoviridae lineage.

3.4.2 | Partiti-Picobirna

35 ant-associated VGF were within the Partiti-Picobirna viral lineage across 15 ant species (Figure 2m). Partiti-Picobirna viruses primarily infect fungi, plants and insects. Seven ant-associated viruses fell outside all viruses within this Partiti-Picobirna phylogeny, which included: Labidus coecus Partiti-Picobirna VGF-1, -3, Ectatomma tuberculatum Partiti-Picobirna VGF-2, Cephalotes atratus Partiti-Picobirna VGF-1, -2, and Crematogaster ant Partiti-Picobirna VGF-4, and Solenopsis virulens Partiti-Picobirna VGF-1. This clade possibly represents an ancient divergence of ant-associated viruses from Partiti-Picobirna viruses. Within the Partitiviridae lineage, nine ant-associated VGF clustered within their own distinct ant-associated clade. These VGF included Gigantiops destructor Partiti-Picobirna VGF-1, -2, Dolichoderus bispinosus Partiti-Picobirna VGF-3, Dolichoderus attelaboides Partiti-Picobirna VGF-1, Odontomachus hastatus Partiti-Picobirna VGF-1, Crematogaster ant Partiti-Picobirna VGF-1, -3 and Ectatomma tuberculatum Partiti-Picobirna VGF-3, -5. Additionally, there were 19 other ant-associated VGF within phylogeny.

3.4.3 | Reo

Reovirales have a widespread host range across plants, invertebrates and vertebrates. Both of the two ant-associated VGF

within the Reovirales were found in the ant species *Odontomachus haematodus* within the same colony (Figure 2q). Odontomachus haematodus Reo VGF-2 clustered within the Orthoreovirus genus and Odontomachus haematodus Reo VGF-1 was sister to this clade.

3.4.4 | Weivirus

There were six ant-associated VGF within the relatively newly described Weivirus clade (Figure 2t). Weiviruses have been found to be associated with mostly arthropod hosts. Within this phylogeny, Crematogaster ant Weivirus VGF-1, -2 and -3 clustered together, and Dolichoderus bispinosus Weivirus VGF-1, Ectatomma tuberculatum Weivirus VGF-1, and Cephalotes atratus Weivirus VGF-1 all were recovered within the phylogeny.

3.5 | Novel dsDNA viruses

3.5.1 | Caudoviricetes

There are over 3500 viral species within Caudoviricetes, accounting for over 30% of all classified viruses. Within this Caudoviricetes bacteriophage phylogeny, there were 57 ant-associated Caudoviricetes PGF across 17 ant species and one complete Caudoviricetes bacteriophage genome within *Dolichoderus bispinosus* (Figure 2d). The ant-associated bacteriophages were recovered quite evenly across the entire Caudoviricetes phylogeny, not clustering together within a specific bacteriophage lineage.

3.5.2 | Lefavirales

Within the Lefavirales phylogeny there were two ant-associated VGF, within *Gigantiops destructor* and *Dolichoderus bispinosus* and one complete Lefavirales virus within *Ectatomma brunneum* (Figure 2h). Lefavirales viruses commonly infect arthropods and are frequently symptomatic within infected individuals. Dolichoderus bispinosus Lefavirales VGF-2 and Ectatomma brunneum virus 1 fell within the Nudiviridae lineage. Nudiviruses often cause disease within their insect hosts including death in the larvae and chronic disease within adults. Gigantiops destructor Lefavirales VGF-4 was most sister to both Musca domestica salivary gland hypertrophy virus and Glossina pallidipes salivary gland hypertrophy virus which are both within *Hytrosaviridae*. These viruses hinder both fertility and fecundity within *Musca domestica* and *Glossina pallidipes*, and the discovery of a salivary gland hypertrophy virus within an ant opens the possibility of *Hytrosaviridae* viral infection within hymenopterans.

3.6 | Novel ssDNA viruses

3.6.1 | Anelloviridae

One complete ant-associated virus fell within the Anelloviridae viral family: Dolichoderus bispinosus Anelloviridae virus 1 and was most closely related to Gyrovirus-11 (Figure 2b).

3.6.2 | Cress

We identified 637 VGF and 181 complete circular genomes of ssDNA CRESS viruses across 30 ant species (Figure 2e). 166 of these complete genomes are newly discovered, with others found within previously described dragonfly and bat-associated viral clades (Table S2). A hallmark of all Rep proteins of CRESS viruses are the conserved RCR-endonuclease and SF3 helicase domains (Rosario, Duffy, & Breitbart, 2012). We were able to identify RCR, SF3-helicase and nonanucleotide motifs for 181 of 637 Rep-encoding CRESS viruses from this study. The number of complete CRESS genomes does not correlate with number of total ant samples from each subfamily (Figure S4). The majority of the complete ant-associated CRESS viruses identified fell outside of described CRESS viral families (137 complete CRESS viruses and 562 VGF).

However, several newly identified ant-associated CRESS viruses clustered within the CRESS viral families Circoviridae (35 complete CRESS viruses including 11 viral isolates and 25 VGF) and Genomoviridae (5 complete viruses including 4 viral isolates and 42 VGF) clades (Figures S5 and S6 illustrate the complete CRESS viral phylogenies). Within Smacoviridae, there were five VGF (Dolichoderus bispinosus CRESS VGF-54, Labidus coecus VGF-1, Ectatomma tuberculatum CRESS VGF-22 and -54, and Odontomachus haematodus VGF-9) and across Geminiviridae – three ant-associated VGF: Crematogaster ant CRESS VGF-5, Odontomachus bauri CRESS VGF-2 and Dolichoderus bispinosus CRESS VGF-2.

3.6.3 | Cruciviridae

There were 35 ant-associated VGF and 15 ant-associated complete genomes within the Cruciviridae phylogeny and these viruses were evenly spread across the Cruciviridae phylogeny (Figure 2f).

3.6.4 | Parvoviridae

There were 166 ant-associated VGF and 53 ant-associated complete viruses within the Parvoviridae phylogeny, across 22 different ant species (Figure 2n). Within Parvovirinae, there were 27 unclassified ant-associated Parvovirinae VGF. One VGF fell within

the Bocaparvovirus–Dolichoderus bispinosus Parvoviridae VGF-2. Within the Brevihamaparvovirus lineage, nine ant-associated viruses clustered including four complete viruses and five VGF. One complete virus was most closely related to the Dependoparvovirus lineage – Crematogaster ant Parvovirus 12. Four VGF clustered within the Chapparvovirus genus. 12 VGF and one complete virus (Dolichoderus bispinosus Parvovirus 6) fell within the Ichthamaparvovirus. Within Densovirinae, there were 61 ant-associated VGF which were most closely related to unclassified Densovirinae. 92 ant-associated VGF and complete viruses fell within Ambidensovirus. Dolichoderus bispinosus Parvoviridae VGF-31 was most closely related to Aquambidensovirus. Two ant-associated VGF fell within the Blattambidensovirus (Dolichoderus bispinosus Parvoviridae VGF-4 and -12). One VGF (Dolichoderus bispinosus Parvoviridae VGF-18) and two complete viruses (Ectatomma tuberculatum Parvovirus 3 and Crematogaster levior Parvovirus 2) clustered within Iteradensovirus. Additionally, five ant-associated VGF fell outside of any classified Parvoviridae clades.

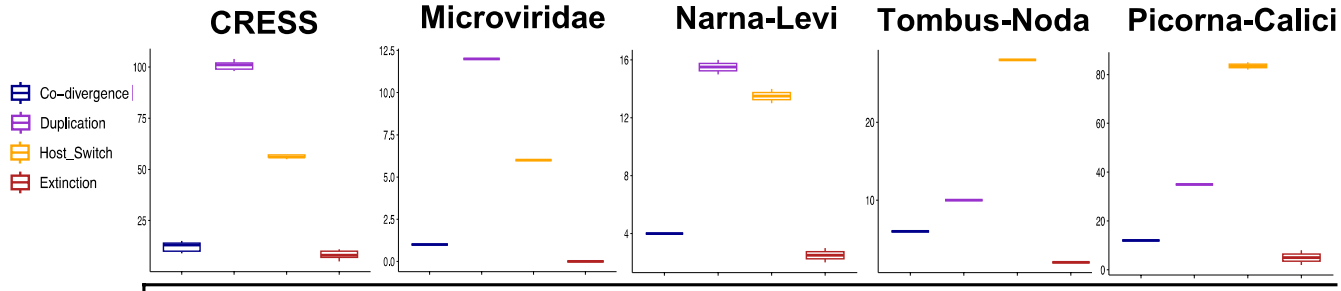
3.6.5 | Microviridae

Overall, we recovered 934 PGF and 405 complete bacteriophage which were associated with bacteria found within the sampled ants (Figure 2j). Currently, there are two subfamilies within Microviridae: Gokushovirinae and Bullavirinae. Bullavirinae bacteriophage tend to infect Enterobacteria, which are known to be facultative symbionts with certain insect species (Moran et al., 2005). Only one ant-associated PGF was found within the Bullavirinae-Dolichoderus bispinosus Microviridae PGF-209. Gokushovirinae are primarily infected by Spiroplasma, Chlamydia, Bdellovibrio and more recently identified in free-living bacteria (Kirchberger & Ochman, 2020). Ant-associated PGF recovered from Gokushovirinae included 385 complete bacteriophage and 856 PGF. 20 complete ant-associated bacteriophage and 77 PGF were most closely related to clades of unclassified Microviridae bacteriophages.

3.7 | Phylogenetic and ecological analyses of viral-host structure

To test for phylogenetic signal between the ant hosts and their associated viruses, we performed three analyses of phylogenetic co-divergence using BaTS, PACo and Jane. Five viral clades were statistically significant across all three analyses. These consisted of two DNA viral clades: CRESS and Microviridae, and three viral RNA clades: Narna-Levi, Tombus-Noda and Picorna-Calici (Figure 4). We visualized the Jane estimates of co-divergence events across the five viral clades which exhibited significant ant-viral phylogenetic association (Figure 4). Overall, these five clades tended to have fewer instances of host switching compared to viral clades which were not found to be significant (Table S3). The tanglegram illustrates the frequent host switching between the ant-associated CRESS virus and the ant host phylogeny (Figure S7).

Four ecological traits were tested (habitat type, nest type, diet type and bacterial abundance of the colony) for association between



	CRESS	Microviridae	Narna-Levi	Tombus-Noda	Picorna-Calici
Jane	x̄=274.28 *p*<0.010	x̄=167.3 *p*<0.010	x̄= 48.7 *p*=0.046	x̄= 70.97 *p*=0.050	x̄=212.28 *p*=0.047
BaTS	AI=0.92; *p*<0.001	AI=0.9; *p*=0.020	AI=0.77; *p*=0.046	AI=0.86; *p*=0.021	AI=0.79; *p*<0.001
PACO	m^2=3992443; *p*=0.050	m^2= 455665; *p*=0.02	m^2= 351130; *p*<0.001	m^2=860429; *p*<0.001	m^2=25435813 ; *p*=0.010

FIGURE 4 Boxplots of estimation of co-phylogenetic events. These co-phylogenetic events are represented across the history of the five ant-associated viral clades which were significant for all three test of host–viral association and co-divergence (CRESS, Microviridae, Narna-Levi, Tombus-Noda and Picorna-Calici). Number of occurrences of different co-evolutionary scenarios for boxplots produced by JANE. Boxplot illustrates the estimated median (centre line), upper and lower quartiles (box limits), and 1.5×interquartile range (whiskers) of the co-divergence (blue), duplication (purple), host-switching (yellow) and extinction (red) events. Table below boxplot represents the results and p-value for each test of co-divergence performed (Jane, BaTS and PACo). For the BaTS AI ratio, A ratio closer to 0 suggests stronger host structure and closer to 1 a weaker host structure.

ant species trait and each viral phylogeny. Two metrics were calculated: δ value for ecological traits within the ant hosts as a measure of phylogenetic signal as well as BaTS to measure the association between ecological traits and the viral clade evolutionary history. We found that across all viruses, four viral clades (Parvoviridae, CRESS, Microviridae and Picorna-Calici) had a statistically significant association for both BaTS and δ between habitat type and the viral phylogeny, with rainforest habitat constituting a significantly larger portion across these viral clades (Tables S4 and S5). Microviridae, Caudoviricetes and Picorna-Calici phylogenies were statistically correlated with the bacterial abundance of the ant host. Only the Picorna-Calici clade had a significant association with diet type and no viral clades had significant association with nest type (Tables S4 and S5).

4 | DISCUSSION

We discovered a staggering diversity of unique RNA and DNA viruses within ants. With the addition of our new ant-associated viruses, our study constitutes 95.8% of all currently described ant viruses, significantly increasing our knowledge of the ant virosphere. The viruses identified were predominantly novel and often divergent from previously known viruses, with over half exhibiting less than 50% AA identity to their top viral protein hit. Similarly to the bee virome (Deboutte et al., 2020), the newly discovered viruses were diverse and often unclassified. Unexpectedly, we were able to identify both DNA and RNA viruses within all ant colonies. Over half of the recovered viruses were ssDNA viruses. This could potentially be due to our methodology, which may preferentially amplify ssDNA. Still, over a third of the recovered VGFs were from RNA viruses and large DNA viruses indicating that these are likely in high abundance within the sampled ant colonies.

This study found that ants are infected by at least 13 different RNA viral clades (Figure 1). Metagenomic data allowed us to characterize 1430 viruses that contained an RNA-dependent RNA polymerase (RdRp) domain, the only universal gene among RNA viruses. Sequence alignments and structural comparisons revealed extensive sequence divergence within these newly discovered RdRp domains, with most sharing less than 50% AA identity with those RNA viruses described previously. The Picorna–Calici clade consists of the large and diverse Picornaviridae and the vertebrate-infecting Caliciviridae, as well as a wide variety of other unclassified viruses. The largest number of ant-associated VGFs identified within this study (1139) belonged in this clade. This abundance is consistent with previous studies (e.g. all *Solenopsis invicta* viruses; Olendraite et al., 2017). Moreover, viruses throughout the clade commonly infect arthropod lineages (Shi et al., 2016).

Mycoviruses are RNA viruses which are known to infect fungus. Six mycoviral genome fragments were identified in *Atta cephalotes* and five in *Labidus coecus*, found within the dsRNA Toti-Chryso viral lineage (Figure 2). Fungus-farming attine ant species (*Atta* and other Myrmicinae genera) obligately depend on the cultivation of specific clades of fungus for food. This close relationship between fungal symbionts and their ant species could explain viral host switching from traditionally fungal-associated viruses to ants. This may be a demonstration of horizontal virus transfer (HVT), the transmission of viruses between often distantly related hosts and is thought to be an important aspect of viral evolution (Dolja & Koonin, 2018). Invertebrates are known to exhibit frequent HVT, often sharing the same viral species with distantly related organisms even in other kingdoms of life (Wolf et al., 2020). *Labidus coecus* is a general predator often preying on the brood of other ants, including fungus-growing ants, like *Atta*. More extensive sampling of fungus-gardening ant lineages would be necessary to better understand the relationship between fungus, ants and mycoviruses since another possibility is that these fungal-associated viruses could be infecting fungi living in the gut of the ant or acquired ephemerally from their diet. To be certain that fungal-associated viruses are truly replicating in the ant host, we would need to perform small RNA sequencing to examine the cellular RNA interference response (Viljakainen et al., 2023). Therefore, further studies are needed to determine active replication and transmission mechanisms of these viruses infecting ant hosts.

Another primarily fungal-infecting clade is the Narna-Levi viruses. Surprisingly, the only complete RNA virus we discovered was within the Narna-Levi clade and identified in a *Cephalotes atratus* ant host. Narna-Levi viruses contain the simplest genomes of any RNA virus and the RdRp domain is highly divergent from other RNA viruses (Hillman & Cai, 2013; Retallack et al., 2021). Though, fungi are the natural hosts of Narna-Levi viruses, they are thought to occasionally infect protists, and now ants.

The DNA viral clades Microviridae, Parvoviridae, and CRESS viruses were the most abundant and well-assembled viruses in our dataset. This was somewhat unexpected since DNA viruses infecting ants were extremely rare until this study. The viral Anelloviridae family consists of small ssDNA viral genomes, commonly found in vertebrates. Within our study, a complete Gyrovirus genome was characterized in a *Dolichoderus bispinosus* ant host, the first time any Anelloviridae virus has been discovered in an invertebrate. The most closely related virus to *Dolichoduerus bispinosus* Gyrovirus 1 was Gyrovirus 11 (68.85% genome-wide pairwise identity). Gyrovirus 11 was discovered in the Ferruginous-backed Antbird (*Myrmeciza ferruginea*), whose diet consists primarily of insects (Mestre et al., 2010). Intriguingly, this bird-associated gyrovirus was found at the exact same isolated Amazonian rainforest locality within Nouragues Reserve, French Guiana. This provides possible evidence of a viral host switching event between distantly related predator and prey species within the same habitat or could be the signal of the bird eating a virus infected insect.

Bacteriophages composed 38.38% of our ant-associated viral dataset and were found across every ant sampled, the most abundant complete viral genomes being Microviridae, small ssDNA viruses. Additionally, Caudoviricetes are a group of large dsDNA viruses, and make up over 30% of all recognized viral species (Walker et al., 2020). The ubiquity of bacteriophage within many ant species is logical due to

the notable associations many bacterial clades exhibit within ant guts (Anderson et al., 2012; Moreau, 2020; Russell et al., 2009). Bee and fly-associated bacteriophages have been widely surveyed and were complex, diverse, and abundant (Deboutte et al., 2020; Kraberger et al., 2019). Not surprisingly, these insect bacteriophages were often the most closely related bacteriophage to the ant-associated bacteriophages within this study. Still, only 10 of 1474 bacteriophage VGFs had >95% AA sequence similarity to their top viral protein hit, implying most of these bacteriophages are divergent from previously known isolates. Though we cannot definitively link the bacteriophage with their bacterial hosts through metagenomic data alone, this study demonstrates their diversity and abundance.

CRESS viruses are circular replication-encoding single-strand DNA viruses within the viral phylum Cressdnaviricota, consisting of seven families. CRESS viruses are widespread across the tree of life (Krupovic et al., 2020). Recent evidence points to cross-kingdom transmission between plant-infecting CRESS viruses and their insect hosts (Czosnek et al., 2017; Liu et al., 2013). For example, a novel CRESS virus was documented in all three obligately symbiotic African ant species that colonize the Whistling-Thorn Acacia tree *Vachellia drepanolobium* (Rosario, Mettel, et al., 2018). This potential interkingdom transmission of CRESS viruses reveal that they may serve a previously unexpected but important role in the ecology of ant ecosystems. Most information on CRESS viruses derive from surveys of environmental samples, like sewage, making precise eukaryotic host-viral linkages impossible to decipher (Kaján et al., 2020). However, since our study pooled ant hosts separately by ant colony, viral-host linkage information is available for co-phylogenetic and ecological trait analyses and revealed 637 viral CRESS genome fragments (181 complete genomes) across 25 ant species. The majority of the ant-associated CRESS viruses identified in this study fell outside the established CRESS viral families, suggesting ant-associated CRESS viral diversity has been grossly underestimated. Many of these ant-associated CRESS viruses form their own ant-specific or insect-specific clades, suggesting they are divergent enough to constitute separate lineages. There were 35 complete CRESS genomes which fell within the Circoviridae family. Of these, nine ant-associated CRESS viruses, all from different ant species hosts were isolates of Dragonfly cyclovirus 5, as all had over 80% genome-wide pairwise identity to this previously identified virus found in *Erythrodiplax umbrata* from Puerto Rico (Table S2). These isolates were derived from varied ant species in geographically distant habitats across French Guiana, indicating geographical expansion of this virus. We cannot definitively explain viral transmission events between dragonfly and ant hosts or if there were potentially intermediate hosts. However, as this nearly identical virus was discovered within so many species of ant from distant locations, it implies there is frequent host switching between ants as well as potentially insect–insect (dragonfly–ant) transmission. Correspondingly, five complete ant-associated CRESS genomes clustered within the Genomoviridae CRESS family from five different ant species (Table S2). Four of these newly identified ant-associated CRESS viruses are isolates of a previously described gemycircularvirus, all of which infect insectivorous bat species (Table S2; Data S1). Until this study, there have been no previous reports of possible ant–predator viral shifts. The discovery of these viral isolates from a bat-associated virus within four separate ant species is compelling because it provides evidence to support viral transmission between ants and insectivorous mammals.

Our data reveal that virus phylogenetic history can mirror that of their ant hosts over long evolutionary timescales. This is supported by the evidence that the CRESS, Microviridae, Tombus-Noda, Narna-Levi and Picorna-Calici clades exhibited significant clustering by host taxonomy, as well as for all host genera and subfamily comparisons. Our co-evolutionary analyses supported congruence and overall host specificity between the viral and ant phylogenies, suggesting that the evolution of these viral clades have generally tracked ant evolution over 150 mya (Moreau et al., 2006). This result is supported by a recent study demonstrating host specificity within ant viruses through siRNA sequencing (Viljakainen et al., 2023). However, despite this overall congruence by ant host species, our data reveal many examples of host-switching throughout virus evolution (Figures 4 and S7; Table S3). Host switching and HVT between ant-associated viruses and distantly related organisms appear to be commonplace throughout their evolutionary history. Host switching is often more frequent than co-divergence across every viral phylogeny (Table S3). Host switching is also indicated throughout ant-associated viruses due to the observation that single viruses are occasionally associated with multiple host species (e.g. in CRESS: Dragonfly cyclovirus 5 isolates). Additionally, unexpectedly over 90% of the clusters of identical viral genomic sequences between ant samples were in different ant subfamilies, indicating ant subfamily-wide host switching of specific viral species (Data S1). Collectively, these results suggest that there is a long-term association between viruses and their ant hosts that stretch back millions of years, but that cross-species transmission has occurred frequently over this background of co-evolution.

We found that the ecological traits of the ant host may play an essential role in the evolution of their viral communities. Of the four ecological traits tested against the viral phylogenies, ant habitat type was significantly correlated within Parvoviridae, Picorna-Calici, Microviridae and CRESS viruses. This means that these viral clades have clustered phylogenetically by whether the colony was found in an urban or rainforest environment, implying the importance of habitat degradation in structuring viral evolution in ants. Habitat loss has the potential to drastically alter the transmission, maintenance, and circulation of viruses as well as the viral community structure within a wild species (Tirera et al., 2021). Additionally, the abundance of bacterial communities within ant colonies may also impact the evolution of Microviridae, Caudoviricetes and Picorna-Calici viruses throughout ant species. Since both Microviridae and Caudoviricetes are bacteriophage lineages, it follows that these clades are associated with bacterial abundance within the ant colony, and could play a critical role in ant evolution. Several studies indicate that symbiotic

bacteria are linked to the evolutionary success of herbivory across ant species, which follows that bacteriophage infecting these symbiotic bacteria are directly impacting the structure of the microbiome in ant species (Russell et al., 2009; Wernegreen et al., 2009).

The 3841 viruses identified in this study fill major gaps in the RNA and DNA virus phylogenies and uncover an evolutionary history that is categorized by both host switching and co-divergence between ants and their associated viruses. These newly discovered VGFs fall across 25 described viral clades and represent a major expansion in the previously known diversity of ant-associated viruses, revealing the remarkable ability of these viruses to switch ant host species over their evolutionary history. Furthermore, frequent horizontal viral transfer between distantly related hosts is suggested, as seen in the ant-associated viruses discovered in the mycoviral lineages (ant-fungus), Gyrovirus 11 and Genomoviridae (predatory–prey associations), and Circoviridae isolates (insect–insect associations). Ecological traits of their ant host species, specifically degradation of habitat, as well as the function of host-associated bacterial communities within the colony may play a role in structuring these viral communities. Further studies should aim to determine the phenotypic consequences of viruses which actively infect ant hosts as well as if there is potential for viral spillover into other hosts if ants experience range expansion due to future habitat disturbance. These findings underscore the critical importance of investigating insect-associated viruses within their natural ecosystems. Such studies are essential for gaining a comprehensive and effective understanding of the evolutionary dynamics between eukaryotic hosts and their microbial communities.

AUTHOR CONTRIBUTIONS

PJF and CSM conceived, designed and executed the study and revised the manuscript. PJF analysed the data and wrote the first draft of the manuscript. PJF and CSM revised and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Anaïs Chanson, Christophe Duplais and Axel Touchard for their indispensable help with fieldwork and to Arielle Salmier, Sourakhata Tirera, Anne Lavergne at the Pasteur Institute of French Guiana. We also thank Maureen Coleman and Cathy Pfister for their helpful discussions and the Grainger Bioinformatics Center for computational support. We extend our sincere thanks to the reviewers for their invaluable feedback and insightful comments, which greatly contributed to the improvement of this manuscript.

FUNDING INFORMATION

PJF was supported by the NSF Graduate Research Fellowship and a Grainger Bioinformatics Center Award. CSM was supported by the National Science Foundation (NSF DEB 1900357 and NSF DBI 2210800).

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

Raw read and nucleotide sequences are available at NCBI under Bioproject: PRJNA1024926. The bioinformatics workflow and code can be found at: https://github.com/peterjflynn/Ant_Associated_Viral_Metagenomics. Trees and alignments for every viral clade can be found at https://doi.org/10.5281/zenodo.11643759.

BENEFITS SHARING STATEMENT

The work provides the benefit of public access to a new viral metagenomic data set obtained from ant colonies across different habitats within French Guiana. Additionally, the workflow and code for all of our analyses are public available on https://github.com/peterjflynn/Ant_Associated_Viral_Metagenomics.

ORCID

Peter J. Flynn https://orcid.org/0000-0002-3182-7671
Corrie S. Moreau https://orcid.org/0000-0003-1139-5792

REFERENCES

Allander, T., Emerson, S. U., Engle, R. E., Purcell, R. H., & Bukh, J. (2001). A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11609–11614. https://doi.org/10.1073/pnas.211424698

Anderson, K. E., Russell, J. A., Moreau, C. S., Kautz, S., Sullam, K. E., Hu, Y., Basinger, U., Mott, B. M., Buck, N., & Wheeler, D. E. (2012). Highly similar microbial communities are shared among related and trophically similar ant species. *Molecular Ecology*, *21*(9), 2282–2296. https://doi.org/10.1111/j.1365-294X.2011.05464.x

Balbuena, J. A., Míguez-Lozano, R., & Blasco-Costa, I. (2013). PACo: A novel procrustes application to cophylogenetic analysis. *PLoS One*, *8*(4), e61048. https://doi.org/10.1371/journal.pone.0061048

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Baty, J. W., Bulgarella, M., Dobelmann, J., Felden, A., & Lester, P. J. (2020). Viruses and their effects in ants (Hymenoptera: Formicidae). *Myrmecological News*, *30*, 213–228. https://doi.org/10.25849/myrmecol.news_030:213

Berthet, N., Reinhardt, A. K., Leclercq, I., Van Ooyen, S., Batéjat, C., Dickinson, P., Stamboliyska, R., Old, I. G., Kong, K. A., Dacheux, L., Bourhy, H., Kennedy, G. C., Korfhage, C., Cole, S. T., & Manuguerra, J. C. (2008). Phi29 polymerase based random amplification of viral RNA as an alternative to random RT-PCR. *BMC Molecular Biology*, *9*, 77. https://doi.org/10.1186/1471-2199-9-77

Bisch, G., Neuvonen, M. M., Pierce, N. E., Russell, J. A., Koga, R., Sanders, J. G., Łukasik, P., & Andersson, S. G. E. (2018). Genome evolution of Bartonellaceae symbionts of ants at the opposite ends of the trophic scale. *Genome Biology and Evolution*, *10*(7), 1687–1704. https://doi.org/10.1093/gbe/evy126

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borges, R., Machado, J. P., Gomes, C., Rocha, A. P., & Antunes, A. (2019). Measuring phylogenetic signal between categorical traits and

phylogenies. *Bioinformatics*, *35*(11), 1862–1869. https://doi.org/10.1093/bioinformatics/bty800

Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59–60. https://doi.org/10.1038/nmeth.3176

Chanson, A., Moreau, C. S., & Duplais, C. (2023). Impact of nesting mode, diet, and taxonomy in structuring the associated microbial communities of Amazonian ants. *Diversity*, *15*(2), 126. https://doi.org/10.3390/d15020126

Conow, C., Fielder, D., Ovadia, Y., & Libeskind-Hadas, R. (2010). Jane: A new tool for the cophylogeny reconstruction problem. *Algorithms for Molecular Biology*, *5*, 16. https://doi.org/10.1186/1748-7188-5-16

Cook, S. C., & Davidson, D. W. (2006). Nutritional and functional biology of exudate-feeding ants. *Entomologia Experimentalis et Applicata*, *118*(1), 1–10. https://doi.org/10.1111/j.1570-7458.2006.00374.x

Czosnek, H., Hariton-Shalev, A., Sobol, I., Gorovits, R., & Ghanim, M. (2017). The incredible journey of Begomoviruses in their whitefly vector. *Viruses*, *9*(10), 273. https://doi.org/10.3390/v9100273

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics*, *27*(8), 1164–1165. https://doi.org/10.1093/bioinformatics/btr088

Davidson, D. W., Cook, S. C., Snelling, R. R., & Chua, T. H. (2003). Explaining the abundance of ants in lowland tropical rainforest canopies. *Science*, *300*(5621), 969–972. https://doi.org/10.1126/science.1082074

Deboutte, W., Beller, L., Yinda, C. K., Maes, P., de Graaf, D. C., & Matthijnssens, J. (2020). Honey-bee–associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(19), 10511–10519. https://doi.org/10.1073/pnas.1921859117

Dolja, V. V., & Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Research*, *244*(August 2017), 36–52. https://doi.org/10.1016/j.virusres.2017.10.020

Duplais, C., Sarou-Kanian, V., Massiot, D., Hassan, A., Perrone, B., Estevez, Y., Wertz, J. T., Martineau, E., Farjon, J., Giraudeau, P., & Moreau, C. S. (2021). Gut bacteria are essential for normal cuticle development in herbivorous turtle ants. *Nature Communications*, *12*(1), 676. https://doi.org/10.1038/s41467-021-21065-y

Feldhaar, H., Straka, J., Krischke, M., Berthold, K., Stoll, S., Mueller, M. J., & Gross, R. (2007). Nutritional upgrading for omnivorous carpenter ants by the endosymbiont Blochmannia. *BMC Biology*, *5*, 48. https://doi.org/10.1186/1741-7007-5-48

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*(23), 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Geoghegan, J. L., Di Giallonardo, F., Wille, M., Ortiz-Baez, A. S., Costa, V. A., Ghaly, T., Mifsud, J. C. O., Turnbull, O. M. H., Bellwood, D. R., Williamson, J. E., & Holmes, E. C. (2021). Virome composition in marine fish revealed by meta-transcriptomics. *Virus Evolution*, *7*(1), veab005. https://doi.org/10.1093/ve/veab005

Gil, R., Silva, F. J., Zientz, E., Delmotte, F., González-Candelas, F., Latorre, A., Rausell, C., Kamerbeek, J., Gadau, J., Hölldobler, B., Van Ham, R. C. H. J., Gross, R., & Moya, A. (2003). The genome sequence of Blochmannia floridanus: Comparative analysis of reduced genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(16), 9388–9393. https://doi.org/10.1073/pnas.1533499100

Guo, J., Bolduc, B., Zayed, A. A., Varsani, A., Dominguez-Huerta, G., Delmont, T. O., Pratama, A. A., Gazitúa, M. C., Vik, D., Sullivan, M. B., & Roux, S. (2021). VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*, *9*(1), 37. https://doi.org/10.1186/s40168-020-00990-y

Hillman, B. I., & Cai, G. (2013). The family Narnaviridae. Simplest of RNA viruses. In *Advances in virus research* (Vol. *86*, 1st ed.). Elsevier Inc. https://doi.org/10.1016/B978-0-12-394315-6.00006-4

Hu, Y., Sanders, J. G., Łukasik, P., D'Amelio, C. L., Millar, J. S., Vann, D. R., Lan, Y., Newton, J. A., Schotanus, M., Kronauer, D. J. C., Pierce, N. E., Moreau, C. S., Wertz, J. T., Engel, P., & Russell, J. A. (2018). Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-recycling gut microbiome. *Nature Communications*, *9*(1), 964. https://doi.org/10.1038/s41467-018-03357-y

Hutchinson, M. C., Cagua, E. F., Balbuena, J. A., Stouffer, D. B., & Poisot, T. (2017). Paco: Implementing procrustean approach to cophylogeny in R. *Methods in Ecology and Evolution*, *8*(8), 932–940. https://doi.org/10.1111/2041-210X.12736

Kaján, G. L., Doszpoly, A., Tarján, Z. L., Vidovszky, M. Z., & Papp, T. (2020). Virus-host coevolution with a focus on animal and human DNA viruses. *Journal of Molecular Evolution*, *88*(1), 41–56. https://doi.org/10.1007/s00239-019-09913-4

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. https://doi.org/10.1093/nar/gkf436

Kirchberger, P. C., & Ochman, H. (2020). Resurrection of a global, metagenomically defined gokushovirus. *eLife*, *9*, e51599. https://doi.org/10.7554/eLife.51599

Kraberger, S., Schmidlin, K., Fontenele, R. S., Walters, M., & Varsani, A. (2019). Unravelling the single-stranded DNA virome of the New Zealand blackfly. *Viruses*, *11*(6), 532. https://doi.org/10.3390/v11060532

Krupovic, M., Varsani, A., Kazlauskas, D., Breitbart, M., Delwart, E., Rosario, K., Yutin, N., Wolf, Y. I., Harrach, B., Zerbini, F. M., Dolja, V. V., Kuhn, J. H., & Koonin, E. V. (2020). Cressdnaviricota: A virus phylum unifying seven families of rep-encoding viruses with single-stranded, circular DNA genomes. *Journal of Virology*, *94*(12), e00582-20. https://doi.org/10.1128/jvi.00582-20

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lester, P. J., Buick, K. H., Baty, J. W., Felden, A., & Haywood, J. (2019). Different bacterial and viral pathogens trigger distinct immune responses in a globally invasive ant. *Scientific Reports*, *9*(1), 5780. https://doi.org/10.1038/s41598-019-41843-5

Letunic, I., & Bork, P. (2007). Interactive tree of life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics*, *23*(1), 127–128. https://doi.org/10.1093/bioinformatics/btl529

Lim, E. S., Zhou, Y., Zhao, G., Bauer, I. K., Droit, L., Ndao, I. M., Warner, B. B., Tarr, P. I., Wang, D., & Holtz, L. R. (2015). Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nature Medicine*, *21*(10), 1228–1234. https://doi.org/10.1038/nm.3950

Liu, B., Preisser, E. L., Chu, D., Pan, H., Xie, W., Wang, S., Wu, Q., Zhou, X., & Zhang, Y. (2013). Multiple forms of vector manipulation by a plant-infecting virus: Bemisia tabaci and tomato yellow leaf curl virus. *Journal of Virology*, *87*(9), 4929–4937. https://doi.org/10.1128/jvi.03571-12

Loope, K. J., Baty, J. W., Lester, P. J., & Wilson Rankin, E. E. (2019). Pathogen shifts in a honeybee predator following the arrival of the varroa mite. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1894), 20182499. https://doi.org/10.1098/rspb.2018.2499

Mestre, L. A. M., Cohn-Haft, M., & Dias, M. M. (2010). Diet and prey availability of terrestrial insectivorous birds prone to extinction in Amazonian forest fragments. *Brazilian Archives of Biology and Technology*, *53*(6), 1371–1381. https://doi.org/10.1590/S1516-89132010000600014

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How many species are there on earth and in the ocean? *PLoS Biology*, *9*(8), e1001127. https://doi.org/10.1371/journal.pbio.1001127

Moran, N. A., Russell, J. A., Koga, R., & Fukatsu, T. (2005). Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Applied and Environmental Microbiology*, *71*(6), 3302–3310. https://doi.org/10.1128/AEM.71.6.3302-3310.2005

Moreau, C. S. (2020). Symbioses among ants and microbes. *Current Opinion in Insect Science*, *39*, 1–5. https://doi.org/10.1016/j.cois.2020.01.002

Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B., & Pierce, N. E. (2006). Phylogeny of the ants: Diversification in the age of angiosperms. *Science*, *312*(5770), 101–104. https://doi.org/10.1126/science.1124891

Nayfach, S., Camargo, A. P., Schulz, F., Eloe-Fadrosh, E., Roux, S., & Kyrpides, N. C. (2021). CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature Biotechnology*, *39*(5), 578–585. https://doi.org/10.1038/s41587-020-00774-7

Nelsen, M. P., Ree, R. H., & Moreau, C. S. (2018). Ant–plant interactions evolved through increasing interdependence. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), 12253–12258. https://doi.org/10.1073/pnas.1719794115

Olendraite, I., Lukhovitskaya, N. I., Porter, S. D., Valles, S. M., & Firth, A. E. (2017). Polycipiviridae: A proposed new family of polycistronic picorna-like RNA viruses. *Journal of General Virology*, *98*(9), 2368–2378. https://doi.org/10.1099/jgv.0.000902

Paez-Espino, D., Eloe-Fadrosh, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N. N., & Kyrpides, N. C. (2016). Uncovering earth's virome. *Nature*, *536*(7617), 425–430. https://doi.org/10.1038/nature19094

Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Parker, J., Rambaut, A., & Pybus, O. G. (2008). Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. *Infection, Genetics and Evolution*, *8*, 239–246. https://doi.org/10.1016/j.meegid.2007.08.001

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Http://Www.R-Project.Org/

Retallack, H., Popova, K. D., Laurie, M. T., Sunshine, S., & DeRisi, J. L. (2021). Persistence of Ambigrammatic Narnaviruses requires translation of the reverse open Reading frame. *Journal of Virology*, *95*(13), e0010921. https://doi.org/10.1128/jvi.00109-21

Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Rice, P., Longden, L., & Bleasby, A. (2000). EMBOSS: The European molecular biology open software suite. *Trends in Genetics*, *16*(6), 276–277. https://doi.org/10.1016/S0168-9525(00)02024-2

Rosario, K., Dayaram, A., Marinov, M., Ware, J., Kraberger, S., Stainton, D., Breitbart, M., & Varsani, A. (2012). Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *Journal of General Virology*, *93*(PART 12), 2668–2681. https://doi.org/10.1099/vir.0.045948-0

Rosario, K., Duffy, S., & Breitbart, M. (2012). A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. *Archives of Virology*, *157*(10), 1851–1871. https://doi.org/10.1007/s00705-012-1391-y

Rosario, K., Fierer, N., Miller, S., Luongo, J., & Breitbart, M. (2018). Diversity of DNA and RNA viruses in indoor air As assessed via metagenomic sequencing. *Environmental Science and Technology*, *52*(3), 1014–1027. https://doi.org/10.1021/acs.est.7b04203

Rosario, K., Mettel, K. A., Benner, B. E., Johnson, R., Scott, C., Yusseff-Vanegas, S. Z., Baker, C. C., Cassill, D. L., Storer, C., Varsani, A., & Breitbart, M. (2018). Virus discovery in all three major lineages of terrestrial arthropods highlights the diversity of single-stranded DNA viruses associated with invertebrates. *PeerJ*, *6*, e5761. https://doi.org/10.7287/peerj.preprints.27093v1

Roux, S., Chan, L. K., Egan, R., Malmstrom, R. R., McMahon, K. D., & Sullivan, M. B. (2017). Ecogenomics of virophages and their giant virus hosts assessed through time series metagenomics. *Nature Communications*, *8*(1), 858. https://doi.org/10.1038/s41467-017-01086-2

Russell, J. A., Moreau, C. S., Goldman-Huertas, B., Fujiwara, M., Lohman, D. J., & Pierce, N. E. (2009). Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(50), 21236–21241. https://doi.org/10.1073/pnas.0907926106

Salmier, A., Tirera, S., De Thoisy, B., Franc, A., Darcissac, E., Donato, D., Bouchier, C., Lacoste, V., & Lavergne, A. (2017). Virome analysis of two sympatric bat species (*Desmodus rotundus* and *Molossus molossus*) in French Guiana. *PLoS One*, *12*(11), e0186943.

Sébastien, A., Lester, P. J., Hall, R. J., Wang, J., Moore, N. E., & Gruber, M. A. M. (2015). Invasive ants carry novel viruses in their new range and form reservoirs for a honeybee pathogen. *Biology Letters*, *11*(9), 20150610. https://doi.org/10.1098/rsbl.2015.0610

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*(7700), 197–202. https://doi.org/10.1038/s41586-018-0012-7

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*(7634), 539–543. https://doi.org/10.1038/nature20167

Simons, A. (2010). A quality control tool for high throughput sequence data.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Steiger, U., Lamparter, H. E., Sandri, C., & Akert, K. (1969). Virus-like particles in the cytoplasm of nervous and glia cells of the ant. Electron microscopical study. *Archiv Fur Die Gesamte Virusforschung*, *26*(3), 271–282.

Sullivan, M. B., Roux, S., Solonenko, N. E., Dang, V. T., Poulos, B. T., Schwenck, S. M., Goldsmith, D. B., Coleman, M. L., & Breitbart, M. (2016). Towards quantitative viromics for both double-stranded and single-stranded DNA viruses. *PeerJ*, *4*, e2777. https://doi.org/10.7717/peerj.2777

Tirera, S., de Thoisy, B., Donato, D., Bouchier, C., Lacoste, V., Franc, A., & Lavergne, A. (2021). The influence of habitat on viral diversity in neotropical rodent hosts. *Viruses*, *13*(9), 1690. https://doi.org/10.3390/v13091690

Valles, S. M., Bell, S., & Firth, A. E. (2014). Solenopsis invicta virus 3: Mapping of structural proteins, ribosomal frameshifting, and similarities to *Acyrthosiphon pisum* virus and kelp fly virus. *PLoS One*, *9*(3), e93497. https://doi.org/10.1371/journal.pone.0093497

Valles, S. M., Porter, S. D., & Calcaterra, L. A. (2018). Prospecting for viral natural enemies of the fire ant Solenopsis invicta in Argentina. *PLoS One*, *13*, e0192377.

Valles, S. M., Strong, C. A., Dang, P. M., Hunter, W. B., Pereira, R. M., Oi, D. H., Shapiro, A. M., & Williams, D. F. (2004). A picorna-like virus from the red imported fire ant, Solenopsis invicta: Initial discovery, genome sequence, and characterization. *Virology*, *328*(1), 151–157. https://doi.org/10.1016/j.virol.2004.07.016

Vibin, J., Chamings, A., Collier, F., Klaassen, M., Nelson, T. M., & Alexandersen, S. (2018). Metagenomics detection and characterisation of viruses in faecal samples from Australian wild birds. *Scientific Reports*, *8*(1), 8686. https://doi.org/10.1038/s41598-018-26851-1

Viljakainen, L., Fürst, M. A., Grasse, A. V., Jurvansuu, J., Oh, J., Tolonen, L., Eder, T., Rattei, T., & Cremer, S. (2023). Antiviral immune response reveals host-specific virus infections in natural ant populations.

Frontiers in Microbiology, *14*, 1119002. https://doi.org/10.3389/fmicb.2023.1119002

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Dempsey, D. M., Dutilh, B. E., Harrach, B., Harrison, R. L., Hendrickson, R. C., Junglen, S., Knowles, N. J., Kropinski, A. M., Krupovic, M., Kuhn, J. H., Nibert, M., Orton, R. J., Rubino, L., Sabanadzovic, S., … Davison, A. J. (2020). Changes to virus taxonomy and the statutes ratified by the international committee on taxonomy of viruses (2020). *Archives of Virology*, *165*(11), 2737–2748. https://doi.org/10.1007/s00705-020-04752-x

Wang, J. H., Xu, G. X., Chen, W., Ma, Y. B., Qi, W., Zhang, C. H., & Cui, X. L. (2021). Impacts of growth form and phylogenetic relatedness on seed germination: A large-scale analysis of a subtropical regional flora. *Ecology and Evolution*, *11*(3), 1280–1293. https://doi.org/10.1002/ece3.7132

Wernegreen, J. J., Kauppinen, S. N., Brady, S. G., & Ward, P. S. (2009). One nutritional symbiosis begat another: Phylogenetic evidence that the ant tribe Camponotini acquired Blochmannia by tending sap-feeding insects. *BMC Evolutionary Biology*, *9*(1), 292. https://doi.org/10.1186/1471-2148-9-292

Wickham, H. (2016). *ggplot2 elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4

Wille, M. (2020). Unravelling virus community ecology in bats through the integration of metagenomics and community ecology. *Molecular Ecology*, *29*(1), 23–25. https://doi.org/10.1111/mec.15306

Wolf, Y. I., Silas, S., Wang, Y., Wu, S., Bocek, M., Kazlauskas, D., Krupovic, M., Fire, A., Dolja, V. V., & Koonin, E. V. (2020). Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. *Nature Microbiology*, *5*(10), 1262–1270. https://doi.org/10.1038/s41564-020-0755-4

Wu, Y., Gao, N., Sun, C., Feng, T., Liu, Q., & Chen, W. H. (2024). A compendium of ruminant gastrointestinal phage genomes revealed a higher proportion of lytic phages than in any other environments. *Microbiome*, *12*(1), 69. https://doi.org/10.1186/s40168-024-01784-2

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, *31*(13), 3406–3415. https://doi.org/10.1093/nar/gkg595

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Flynn, P. J., & Moreau, C. S. (2024). Viral diversity and co-evolutionary dynamics across the ant phylogeny. *Molecular Ecology*, *33*, e17519. https://doi.org/10.1111/mec.17519