Virology 535 (2019) 189–199

Contents lists available at ScienceDirect

Virology

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

Identification of diverse arthropod associated viruses in native Australian fleas

Erin Harvey[a], Karrie Rose[b], John-Sebastian Eden[a,c], Andrea Lawrence[d,e,f], Stephen L. Doggett[g], Edward C. Holmes[a,*]

[a] Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Life and Environmental Sciences and Sydney Medical School, The University of Sydney, Sydney, NSW, 2006, Australia
[b] Australian Registry of Wildlife Health, Taronga Conservation Society Australia, Mosman, NSW, 2088, Australia
[c] Centre for Virus Research, Westmead Institute for Medical Research, Westmead, NSW, 2145, Australia
[d] Medical Entomology, NSW Health Pathology, ICPMR, Westmead Hospital, Westmead, NSW, 2145, Australia
[e] Sydney School of Veterinary Science, The University of Sydney, Sydney, NSW, 2006, Australia
[f] SpeeDx, Pty Ltd., Eveleigh, NSW, 2015, Australia
[g] Department of Medical Entomology, NSWHP-ICPMR, Westmead Hospital, Westmead, NSW, 2145, Australia

ARTICLE INFO

Keywords:
Fleas
Meta-transcriptomics
Virus discovery
RNA virus
Virome
Stephanocircus
Macropsylla
Pygiopsylla

ABSTRACT

Fleas are important vectors of zoonotic disease. However, little is known about the natural diversity and abundance of flea viruses, particularly in the absence of disease associations, nor the evolutionary relationships among those viruses found in different parasitic vector species. Herein, we present the first virome scale study of fleas, based on the meta-transcriptomic analysis of 52 fleas collected along the eastern coast of Australia. Our analysis revealed 18 novel RNA viruses belonging to nine viral families with diverse genome organizations, although the majority (72%) possessed single-stranded positive-sense genomes. Notably, a number of the viruses identified belonged to the same phylogenetic groups as those observed in ticks sampled at the same locations, although none were likely associated with mammalian infection. Overall, we identified high levels of genomic diversity and abundance of viruses in the flea species studied, and established that fleas harbor viruses similar to those seen to other vectors.

1. Introduction

Fleas (Order Siphonaptera) are one of the most common obligate parasites of warm-blooded animals and are important vectors of zoonotic disease (Van der Mescht and Matthee, 2017). While fleas are known to carry a variety of bacterial pathogens including species of *Rickettsia* and *Bartonella* (Kaewmongkol et al., 2011; Lawrence et al., 2015; Schloderer et al., 2006), they are perhaps best known as the primary vector of *Yersinia pestis,* the causative bacterial agent of human plague (Raoult et al., 2013). Pathogens may be carried in the flea midgut and transmitted through regurgitation from the digestive tract when feeding, or contained within flea excrement and transmitted when feces contaminate open wounds (Bitam et al., 2010; Oshima et al., 2016). Although fleas are implicated in the transmission of a number of viral pathogens, including canine distemper virus (CDV) (Trebbien et al., 2014), feline leukemia virus (FeLV) (Vobis et al., 2003), feline calicivirus (Mencke et al., 2009), tick-borne encephalitis virus (Rehacek, 1961) and myxoma virus (MYXV) (Kerr et al., 2015), far less is known about the viruses of fleas than those of other parasitic arthropod vectors, particularly mosquitoes and ticks (Brinkmann et al., 2016, 2018; Pettersson et al., 2017; Shi et al., 2017; Tokarz et al., 2014). Similarly, it is unclear whether most flea viruses are pathogens or commensals.

There are currently some 2574 documented species of flea, belonging to 238 genera and 16 families, although the majority of these do not live in close association with humans (Bitam et al., 2010). In Australia, there are some 88 flea species, 89% of which are endemic (in some cases including neighboring Indonesia and Papua New Guinea) and parasitize native wildlife including marsupials and native rodents (Dunnet and Nardon, 1974). While fleas feed on a wide range of mammal and bird species, ~74% of characterized species are associated with rodents. Fleas are capable of easily moving between host animals, aided by their ability to jump up to 150 times their body length (Bitam et al., 2010). Importantly, due to climate change, the frequency and

* Corresponding author.
E-mail address: edward.holmes@sydney.edu.au (E.C. Holmes).

https://doi.org/10.1016/j.virol.2019.07.010
Received 19 May 2019; Received in revised form 2 July 2019; Accepted 8 July 2019
Available online 09 July 2019
0042-6822/

length of drought periods in grassland areas is rising globally, which in turn is likely to increase the prevalence of fleas and flea-associated infectious disease (Eads and Hoogland, 2017).

Meta-transcriptomics (i.e. bulk shotgun RNA sequencing) has been proven a powerful tool in virus discovery, particularly those with RNA genomes and in the context of arthropod hosts (Shi et al., 2016). Indeed, large-scale meta-transcriptomic studies of invertebrate viruses have greatly changed our understanding of the complexity and structure of the virosphere (Zhang et al., 2018). We have previously used this technique to investigate the virome of ticks sampled in New South Wales, Australia, identifying 19 novel RNA viruses all of which were divergent from previously identified tick associated viruses. Of these novel viruses, three were phylogenetically similar to RNA viruses sampled from mammals suggesting that they may be able to infect mammals. While a number of studies have investigated the virome of ticks in a range of locations globally (Brinkmann et al., 2018; Pettersson et al., 2017; Tokarz et al., 2014), it is unknown if fleas carry similar types and abundance of viruses as other blood feeding parasitic invertebrates.

Herein, we used a meta-transcriptomic approach in the first study of the flea virome. As with our previous study of ticks, fleas were collected from the North shore area of Sydney and Timbillica national park in New South Wales on the central east coast of Australia. All fleas were collected while feeding on native bush rats, black rats or native marsupials. Our study aimed to (i) characterize the total diversity and abundance of viruses carried in a range of native Australian flea species, (ii) determine whether these virus species are similar to those found in Australian ticks collected in the same regions from the same hosts, and (ii) use phylogenetic analysis to help determine if any of these viruses are likely to replicate in mammalian hosts.

2. Materials and methods

2.1. Sample collection and flea identification

Fleas were collected live, placed into a liquid nitrogen dewer for transportation, and frozen at −80 °C during bandicoot population studies on the North shore area of Sydney (33.8224° S, 151.2994° E) and Timbillica State Forest (37.3712° S, 149.7211° E), New South Wales, Australia. Fleas were collected from long-nosed bandicoots (*Perameles nasuta*), long-nosed potoroos (*Potorous tridactylus*), southern brown bandicoots (*Isoodon obesulus*), as well as from invasive black rats (*Rattus rattus*) and native bush rats (*Rattus fuscipes*) unintentionally caught in bandicoot traps. All animals trapped in Timbillica State Forest were handled in accordance with procedures approved by the Australian National University Animal Ethics Committee (A2015/26), while those trapped in Sydney were handled in accordance with methods approved by the New South Wales Office of Environment and Heritage Animal Ethics Committee (000214/05), each in compliance with the Australian Code of Practice for the Use of Animals for Scientific Purposes. Samples were collected under a scientific license provided by the NSW Office of Environment and Heritage (number SL100104).

Fleas were morphologically identified to the species level using keys and descriptions (Dunnet and Nardon, 1974). Identification was performed using a stereomicroscope and a cold table to retain RNA integrity. For some specimens, identification to species level required slide-mounting of flea specimens which was not possible due to subsequent RNA extraction. These fleas were identified to the genus level only.

2.2. Extraction of RNA, pooling and sequencing

Fleas were washed in DPBS solution and homogenized in 800 µl lysis buffer using a TissueRuptor (Qiagen). Total RNA was extracted using a RNeasy plus mini kit (Qiagen) according to manufacturer's instructions. RNA quality was assessed using an Agilent 2100 bioanalyzer (Agilent Technologies). Samples were pooled for sequencing based on the genus of flea and that of the flea's host (marsupial, bush rat, black rat). Libraries were constructed using a TruSeq total RNA library preparation kit (Illumina) and host rRNA was depleted using a Ribo-Zero-Gold (Human, Mouse, Rat) kit (Illumina). Paired-end sequencing was performed on the Hiseq2500 platform (Illumina) at the Australian Genomic Research Facility (AGRF).

2.3. Assembly and analysis

Sequencing reads were trimmed for quality using trimmomatic and *de novo* assembled using Trinity (2.2.1) (Grabherr et al., 2011). The assembled contigs were blasted against the NCBI nr protein database using diamond blastx (v.0.9.10) (Buchfink et al., 2015) and contigs with significant sequence similarity to viral proteins were selected. The selected contigs were then screened through NCBI BlastX for sequence similarity to non-viral sequences and the remaining sequences were searched for predicted ORFs using ExPASy translate tool (https://web.expasy.org/translate/). The predicted ORF structure of the contig was then compared to the genome of the closest blast hit to determine if the contig was potentially an endogenous viral artefact.

2.4. Phylogenetic analysis

To determine the evolutionary history of each virus, the RNA-dependent RNA polymerase (RdRp) or polyprotein of each virus was compared to previously described proteins from the relevant virus family. The RdRp and polyprotein sequences for the background data set of viruses from each virus family group were retrieved from the NCBI RefSeq database (https://www.ncbi.nlm.nih.gov/Taxonomy/Browser). Multiple sequence alignment was performed with Mafft (v.7.300) (Katoh and Standley, 2013) using the L-INS-I algorithm, and all ambiguously aligned were removed using TrimAL (v.1.4.1) (Capella-Gutierrez et al., 2009) with a maximum gap threshold of 0.8 and minimum similarity threshold of 0.005. IQ-TREE (v.1.6.1) (Nguyen et al., 2015) was used to determine the best-fit model of amino acid substitution for each sequence alignment (Table 2). Maximum likelihood phylogenetic trees based on the amino acid alignments were then estimated with the PhyML program (Guindon et al., 2010) employing the best-fit model suggested by IQ-TREE, with 1000 bootstrap replications (although severe computational constraints meant that the expansive phylogeny of the picorna-like viruses could only be run with 100 bootstrap replications).

2.5. Characterization of flea 18S ribosomal (r) RNA sequences and co-infecting bacteria and microbial eukaryotes

The assembled contigs were compared to the NCBI nt database using Blastn and filtered for > 90% sequence similarity over > 200 nt with an e-value of < 1e-25 as an arbitrary cut-off. This subset of blast results was then filtered to identify the bacterial composition using the 16S rRNA gene, while 18S rRNA sequences were used to confirm the species identity of the flea host in each library. Contigs showing sequence similarity to the 18S rRNA gene of flea species were extracted. The selected nucleotide sequences were checked for open reading frames, and the resulting amino acid sequences were used in an NCBI web blast, with the top blast results used to infer phylogenetic trees. The nucleotide sequences of flea species related to those detected were extracted from NCBI and aligned to the contigs extracted from our data using the Mafft alignment tool. The resulting alignment was then used to infer a nucleotide sequence maximum likelihood tree using the procedure described above. Specifically, the ModelFinder function of IQ-TREE was used to determine the appropriate nucleotide substitution model, which was found to be K80 for the flea 18S rRNA alignment. The PhyML package was then used to estimate phylogenetic history, with 1000 bootstrap replications. Co-infecting microbial eukaryotes, including

Table 1
Composition of RNA pools for sequencing and the number of sequencing reads produced for each library.

Library	Flea species or group	Host type	Number of individuals	Number of reads	Sampling location, NSW
1	*Stephanocircus harrisoni*	Long nose bandicoot, Southern brown bandicoot, Potoroo	15	37,847,360	Timbillica
2	*Pygiopsylla* Group A	Long nose bandicoot, Southern brown bandicoot, Potoroo	15	38,613,790	Timbillica, North head, Sydney
3	*Pygiopsylla* Group A, *Pygiopsylla rainbowi*	Native rat	3	35,639,160	Timbillica, North head, Sydney
4	*Macropsylla hercules*	Native rat	6	36,332,036	Timbillica
5	*Stephanocircus pectinipes, Stephanocircus harrisoni*	Native rat	11	35,894,390	Timbillica
6	*Pygiopsylla rainbowi*	Black rat	2	38,338,622	North head, Sydney

fungi, were identified in the assembled contigs by taxonomic ranking using CCMetagen (v. 0.1) (Marcelino et al., 2019), which relies on the sequence alignments produced by the mapping tool KMA (Clausen et al., 2018). The CCMetagen output was then filtered to identify all hits classified within the Eukayota, although those to the phyla Arthropoda and Chordata were removed, as were any to uncharacterized organisms or to cloning vectors.

2.6. Abundance measurements

rRNAs were excluded from the total read count by identifying rRNA contigs in the assembled data using Blastn against the NCBI nt database. All hits to rRNAs with a sequence similarity of over 90% across more than 200 nt and with an e-value of less than 1e-25 were extracted. Bowtie2 (Langmead and Salzberg, 2012) was used to align the fastq reads back to the contigs using end-to-end alignment with no errors allowed. The total number of ribosomal reads was then subtracted from the total read count. Bowtie2 was used to align fastq reads to each novel virus genome or contig. This was then repeated for the host COX1 gene. Bacterial abundance was measured by identifying 16S rRNA from the results of a Blastn search of all assembled contigs against the GenBank nt database. The abundance of contigs showing > 90% nucleotide sequence similarity to a 16S rRNA sequence was then measured using Bowtie2, and the number of reads per kilobase million (RPKM) was calculated for each contig. Similarly, the eukaryotic microbial hits identified using CCMetagen were extracted from the Trinity assembly, and Bowtie2 was used to map the raw reads back to these contigs. The RPKM was then calculated as described above.

2.7. Accession numbers

The virus consensus sequences obtained here have been submitted to GenBank and assigned accession numbers MN167477-MN167503. The raw read data have been submitted to the NCBI Sequence Read Archive (SRA) database and assigned BioProject ID PRJNA542968.

3. Results

3.1. Flea species assignment

A total of 52 fleas were collected during bandicoot population studies in two locations on the east coast of Australia during 2016 and 2017 (Table 1). Fleas were identified as belonging to three genera - *Stephanocircus, Pygiopsylla* and *Macropsylla* - and four species - *Stephanocirus harrisoni, Stephanocirus pectinipes, Pygiopsylla rainbowii* and *Macropsylla hercules*. A fifth group of fleas were identified as belonging to *Pygiopsylla* Group A. Identification to species level was not possible in all cases as fleas often require slide mounting for species determination, which cannot be performed while maintaining temperatures low enough to preserve RNA quality (Dunnet and Nardon, 1974). The RNA from these 52 fleas were pooled by flea genus and the host from which the flea was collected, producing six libraries for RNA sequencing (Table 1). The resulting RNA sequencing produced between 35,639,160 and 38,613,790 paired end reads per library which were *de novo* assembled into between 85,994 and 153,532 contigs per library (Table 1).

Flea 18S rRNA sequences were identified from the assembled contigs and a phylogenetic analysis was used to confirm the species identification. As no reference 18S rRNA sequences exist in the NCBI RefSeq database for the species in this study, the most closely related COX1 sequences were included in the analysis. Sequences from libraries 2, 3 and 6, representing fleas belonging to the genus *Pygiopsylla*, grouped together loosely with other *Pygiopsylla* reference sequences as expected based on morphological classification (Fig. 1, Table 1). In contrast, the sequences from libraries 1 and 5, collected from fleas belonging to the genus *Stephanocircus*, grouped together with the reference *Stephanocircus* sequences, again expected based on the morphological

Table 2
Description of the novel viruses identified in this study.

Virus name	Closest virus family	Length (nt)	Number of reads	Library	Segments	Coding complete segments (Y/N)
Sherlock virus	*Narnaviridae*	2913	2822	1	1	Y
Mycroft virus	*Sobemoviridae*	2717	7658	1	1	Y
Lestrade virus	*Orthomyxoviridae*	10910	79913	2	5	Y
Watson virus	*Picornaviridae*	10080	3617	2	1	Y
Baskerville virus	*Picornaviridae*	8713	2087	2	1	Y
Mortimer virus	*Sobemoviridae*	4924	10757	2	2	Y
Barrymore virus	*Totiviridae*	2012	133	2	1	Y
Stapleton virus	*Sobemoviridae*	4332	6436	3	2	Y
Stamford virus	*Picornaviridae*	7447	38948	3	1	Y
Gregson virus	*Picornaviridae*	5290	35399	3	1	Y
Hudson virus	*Reoviridae*	7071	976	3	2	Y
Moriarty virus	*Narnaviridae*	2227	158	3	1	N
Brunton virus	*Narnaviridae*	2503	177	3	1	N
Carfax virus	*Picornaviridae*	9136	4802	4	1	Y
Culverton virus	*Chuviridae*	15278	33835	4	1	Y
Cushing virus	*Tombusviridae*	2209	26161	5	1	Y
Browner virus	*Phenuiviridae*	8743	5622	5	2	Y
Moran virus	*Picornaviridae*	10634	16273	6	1	Y

classification of the fleas included in these libraries (Fig. 1, Table 1). The 18S rRNA sequence identified in library 4 appears in an divergent group with a single related reference sequence, *Macropsylla novaehollandiae*, again as expected based on the morphological identification (Fig. 1, Table 1).

3.2. RNA virus diversity

A total of 18 novel RNA viruses were identified, comprising 14 complete coding sequences and four partially assembled genomes. No previously described viruses were identified, nor viruses with DNA genomes. The 18 novel RNA viruses showed phylogenetic relationships to nine viral families: *Narnaviridae, Sobemoviridae, Picornaviridae, Totiviridae, Chuviridae, Reoviridae, Orthomyxoviridae, Tombusviridae* and *Phenuiviridae*. All major types of RNA virus genome structure were represented within the data, with those possessing positive-sense single stranded genomes the most common, accounting for 72% of species identified and 56% of the total number of viral reads within the total data set. We also predict that these sequences are not endogenous viral elements as they show no relationship to previously described flea genomic sequences and contain ORFs unbroken by premature stop codons. Finally, the abundance and diversity of viruses does not appear to be associated with the number of fleas contained within the library pool. In particular, libraries 1 and 2 contained the same number of individual fleas, while library 3 contained only 20% of that number, yet libraries 2 and 3 harbored the greatest viral diversity while library 1 possessed just two distinct virus species (Table 1, Table 2).

Host gene expression was measured using the abundance of the COX1 housekeeping gene. Accordingly, the level of expression varied from 0.6 to 1.6% of reads with rRNA sequences excluded (Fig. 2). Similarly, the total abundance of viral reads was assessed for each library as a percentage of the total non-rRNA sequences. This revealed that libraries 2 and 3 exhibited a substantially higher abundance of viral reads (0.28% and 0.26% respectively; Fig. 2). These two libraries also

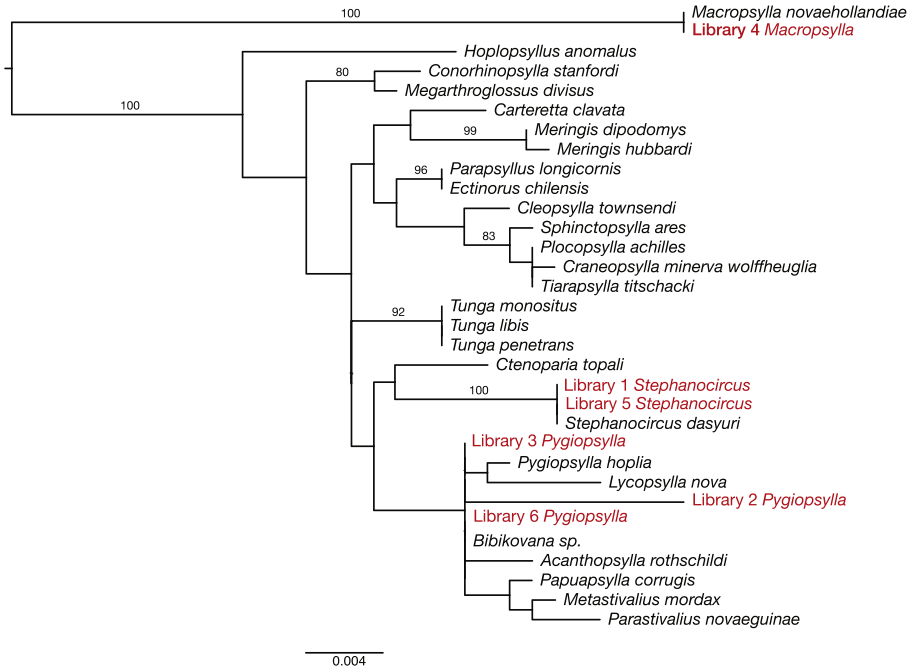

Fig. 1. Maximum likelihood phylogeny of the 18S rRNA sequences of the flea taxa present within each library and reference 18S rRNA sequences of related flea species. Reference sequences are shown in black while sequences generated in this study are shown in red. Bootstrap values of > 70% are shown.

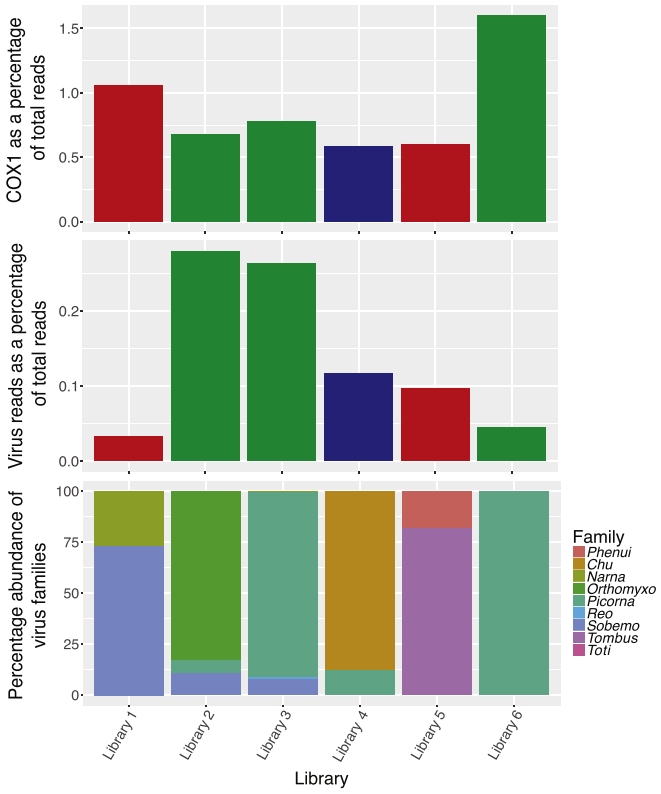

Fig. 2. Abundance of COX1, viral reads and virus families. The abundance of the COX1 housekeeping gene and total viral sequences are shown as a percentage of the total number of non-rRNA sequences. Columns are colored by the flea genera in each library with *Stephanocircus* shown in red, *Pygiopsylla* in green and *Macropsylla* in blue. The breakdown of virus family abundance is represented, with the total number of reads of each virus family shown as a percentage of the total number of viral reads in each library.

contained the greatest diversity of RNA virus species (Table 2) and the two most highly abundant viruses within the data set - the orthomyxo-like Lestrade virus and the picorna-like Stamford virus. Interestingly, the three most abundant viruses within the data set were from different families - the *Orthomyxoviridae*, *Picornaviridae* and *Chuviridae* - although picorna-like viruses were the most abundant family in all three *Pygiopsylla* libraries (Tables 1 and 2, Fig. 2). We now describe each group of viruses in turn.

3.3. Single-stranded, positive-sense RNA viruses

Of the 13 single-stranded (ss), positive sense (+) RNA viruses identified here, six belonged to the *Picornaviridae* and were found in four of the six libraries. Picorna-like viruses were absent from *Stephanocircus* libraries (Table 2, Fig. 3), although whether this merely reflects chance sampling is unclear. Despite clustering within the same family and with other invertebrate associated viruses, the six viruses from the *Picornaviridae* were not closely related to each other. Indeed, the most closely related species were Watson virus and Carfax virus that exhibited only 40% amino acid sequence similarity in the RdRp.

The abundance of the picorna-like viruses varied dramatically between species, with the most abundant, Stamford virus, comprising 0.12% of the total non-rRNA reads within library 3 and the least abundant, Baskerville virus, constituting 0.006% of the total non-rRNA reads in library 2 (Fig. 2). The two most highly abundant picorna-like viruses clustered with the same group within the picorna-like virus tree, unofficially referred to as “noraviruses”, although they exhibit little amino acid sequence similarity (Figs. 2 and 3). These viruses are all arthropod-associated and many viruses within this cluster were identified within *Drosophila* (Fig. 3). Unlike all other picorna-like viruses identified in this data set that contained a single ORF, Stamford virus comprised two ORFs. This genome structure is also seen in viruses related to Stamford virus, such as a *Drosophila*-associated Nora virus. The assembled complete coding genomes of the picorna-like viruses identified here varied from 5320 nt to 10,633 nt (Fig. 3).

The remaining seven + ssRNA viruses fell into three other family groups: three within the *Sobemoviridae,* three within the *Narnaviridae*, and a single virus in the *Tombusviridae.* These viruses were generally of lower abundance than the picorna-like viruses described above, with the most abundant being Mortimer virus which represented 0.03% of reads in library 2, and the lowest being Moriarty virus with 0.0005% of the reads in library 3. Most of the previously identified viruses within the three trees are associated with invertebrates, although a number of viruses within the *Narnaviridae* tree are found in fungi, plants and protists (Fig. 4). Of the three narna-like viruses identified in this data set, a complete coding genome could only be assembled for Sherlock virus, likely because the remaining two viruses (Moriarty virus and Brunton virus) were are at low abundance (Figs. 2 and 4). Phylogenetic analysis of the narna-like viruses revealed that although Moriarty virus and Sherlock virus only show ~23% amino acid sequence similarity they fell into the same phylogenetic group (Fig. 4). This is notable as both Brunton and Moriarty viruses were identified in *Pygiopsylla* fleas collected from native rats, while Sherlock virus was sampled from *Stephanocircus* fleas feeding on marsupials (Table 2). Similarly, Mycroft virus and Stapleton virus group together (48% amino acid sequence similarity) within the sobemo-like tree despite being associated with different flea species. A single species of tombus-like virus, denoted Cushing virus, was also identified.

3.4. Single-stranded, negative-sense RNA viruses

Three novel negative (−) sense ssRNA viruses were identified here, each belonging to a different family: the *Chuviridae*, *Phenuiviridae* and *Orthomyxoviridae*. All three viruses clustered phylogenetically with other invertebrate associated viruses (Fig. 5). Notably, the two most highly abundant virus species identified exhibited this form of genome organization - the orthomyxo-like virus, Lestrade virus (0.23% of non-rRNA reads in library 2), and the Chu-like virus, Culverton virus (0.1% of non-rRNA reads in library 4) (Fig. 2). Five orthomyxo-like segments were identified within library 2, all belonging to Lestrade virus. Two segments of the Phenui-like virus, Browner virus, were also identified, consistent with the other viruses within this group.

3.5. Double-stranded RNA viruses

Double-stranded (ds) RNA viruses were the least represented within the data set with just two virus species of this type - Barrymore virus that fell within the *Totiviridae* and Hudson virus that clustered with the *Reoviridae* (Fig. 6). These viruses were also of relatively low abundance, with Barrymore virus accounting for just 0.0004% of reads in library 2, and Hudson virus accounting for 0.003% of the reads in library 3 (although these two libraries contain the most virus reads and the greatest abundance of virus species). The majority of the reference sequences within both dsRNA trees are invertebrate associated, although a number of the viruses within the *Totiviridae* tree are associated with fungi, protists or plants. Notably, the two virus species most closely related to Barrymore virus are found in fungi, suggesting that this virus may in fact be associated with flea-infecting fungi rather with than the flea itself, although this cannot be determined conclusively from the phylogenetic data alone.

3.6. Co-infecting bacteria, fungi and protozoa

Other microbial species are commonly sampled with viruses within individual invertebrates (Harvey et al., 2019). To address this issue we

Fig. 3. Phylogenetic relationships and genomic structures of the novel picorna-like viruses described here. Phylogenies are based on amino acid sequences of the RdRp region of the polyprotein. A phylogeny of all the picorna-like viruses used in the analysis is shown in the upper-left panel. The novel flea viruses identified here fall into two clades - marked in red in the main tree - that are then analyzed in more detail in the phylogenies on the right. All trees are scaled to the number of amino acid substitutions per site and midpoint rooted for clarity. Bootstrap values of more than 70% are shown. Viruses shown in blue are associated with the genus *Macropsylla*, while those shown in green and red are associated with *Pygiopsylla* and *Stephanocircus* genera, respectively. Branch tips are colored according to virus hosts, with fungi and plant viruses indicated by an orange branch, vertebrate viruses with a purple branch, and invertebrate viruses with a black branch. Genome diagrams provide information on the number of genome segments identified, the length in nucleotides of each segment, the number of predicted ORFs, and any predicted conserved protein structures.

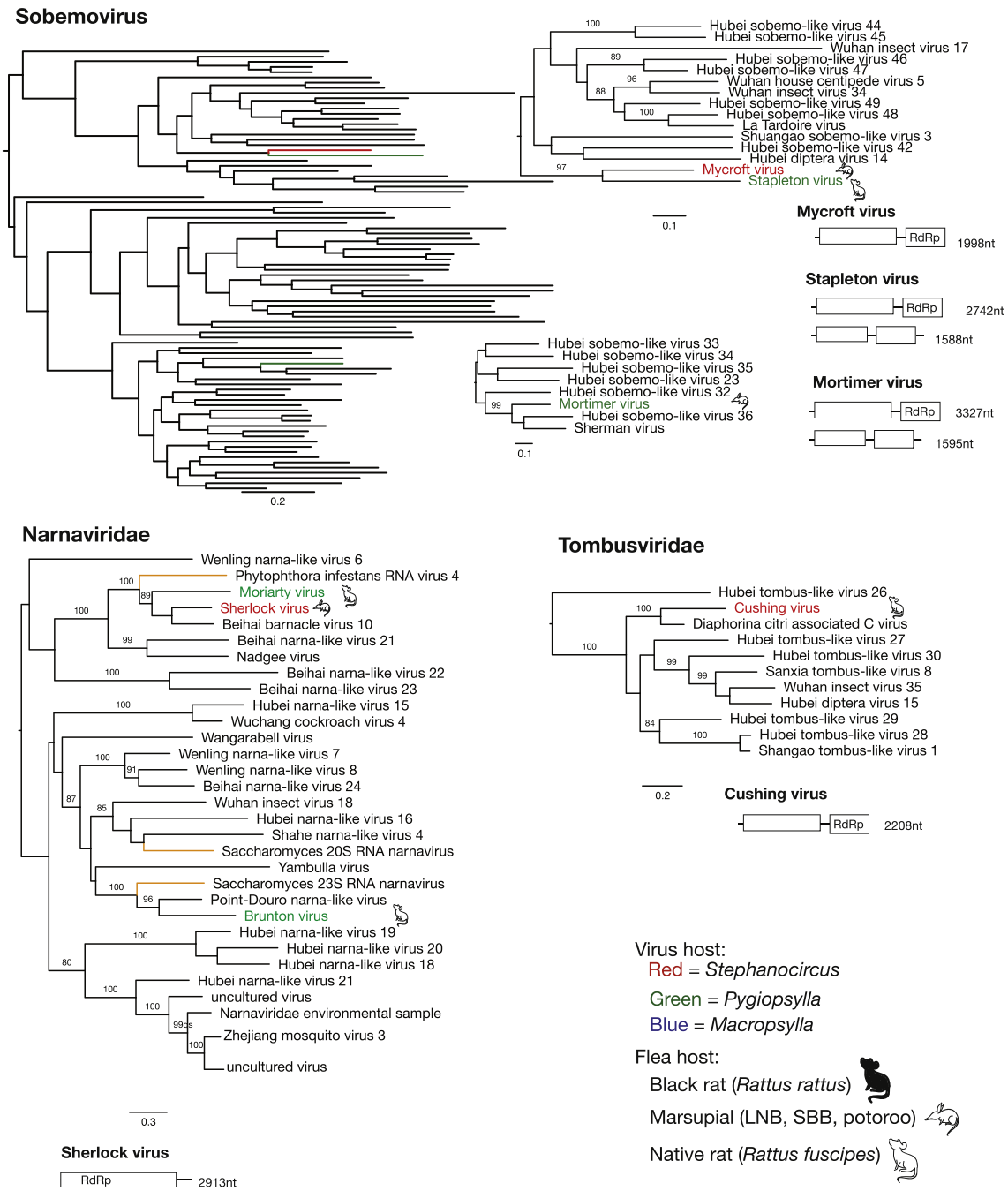

Fig. 4. Phylogenetic relationships of the RdRp region and genomic structures of the novel ssRNA + viruses identified here (excluding picorna-like viruses; see Fig. 3). All trees are scaled to the number of amino acid substitutions per site and midpoint rooted for clarity. Bootstrap values of more than 70% are shown. Viruses shown in blue are associated with the genus *Macropsylla*, while those shown in green and red are associated with *Pygiopsylla* and *Stephanocircus* genera, respectively. Branch tips are colored according to virus hosts, with fungi and plant viruses indicated by an orange branch, vertebrate viruses with a purple branch, and invertebrate viruses with a black branch. Genome diagrams provide information on the number of genome segments identified, the length in nucleotides of each segment, the number of predicted ORFs, and any predicted conserved protein structures.

assessed the microbial composition of each library via the abundance of the 16S/18S rRNA hits in the blastn results for bacteria (Table 3) and using the KMA and CCMetagen tools for eukaryotic microbes (Clausen et al., 2018; Marcelino et al., 2019).

Bacterial hits were grouped by genus, although for a number of the 16S rRNA hits no genus or species were listed in the RefSeq description such that they were grouped as 'uncultured bacteria'. The greatest diversity of bacterial genera was seen in both *Stephanocircus* libraries (libraries 1 and 5), with six and seven genera identified, respectively. The genus *Bacillus* was identified in all libraries other than library 4 (*Macropsylla*). Bacteria of the genus *Bartonella* were identified in libraries 4 and 5 (*Macropsylla* and *Stephanocircus*). Although *Bartonella* have previously been identified in fleas (Oshima et al., 2016), the 16S rRNA contigs exhibited the highest sequence similarity to two rodent associated species, *Bartonella japonica* and *Bartonella rattaustraliani*.

The greatest abundance of protozoa and fungi were seen in libraries 1, 3 and 5, while libraries 1 and 5 had the greatest diversity in terms of the number of genera identified (Table 4). The three genera with the highest abundance in a given library were *Trypanosoma*, *Blechomonas* and *Leptomonas*, all of which are all members of the order

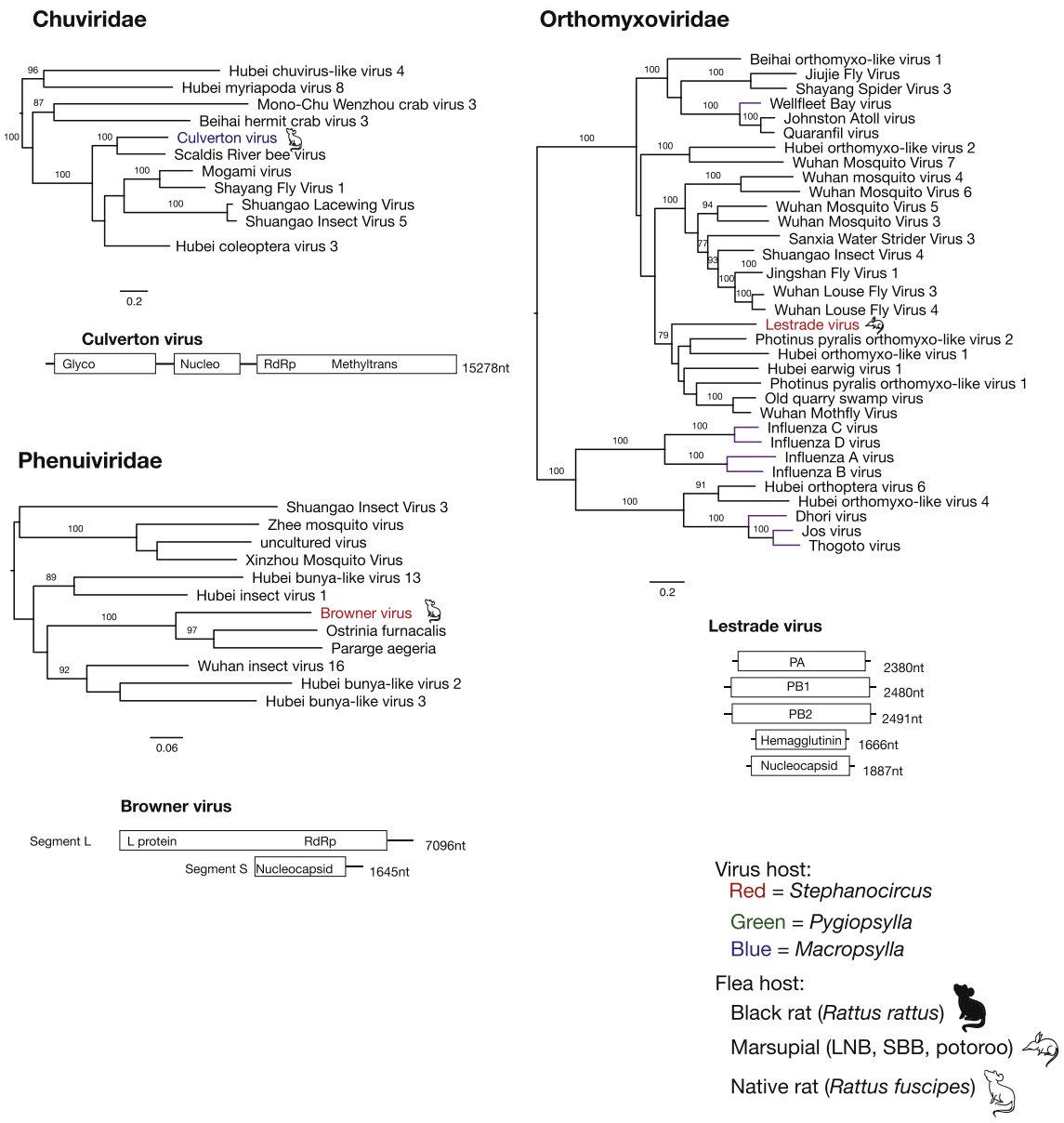

Fig. 5. Phylogenetic relationships and genomic structures of the novel ssRNA-viruses. All trees are scaled to the number of amino acid substitutions per site and midpoint rooted for clarity. Bootstrap values of more than 70% are shown. Viruses shown in blue are associated with the genus *Macropsylla*, while those shown in green and red are associated with *Pygiopsylla* and *Stephanocircus* genera, respectively. Branch tips are colored according to virus hosts, with fungi and plant viruses indicated by an orange branch, vertebrate viruses with a purple branch, and invertebrate viruses with a black branch. Genome diagrams provide information on the number of genome segments identified, the length in nucleotides of each segment, the number of predicted ORFs, and any predicted conserved protein structures.

Trypanosomatida (Table 4). The next most abundant eukaryotic taxon, and the most abundant fungi in the data set, was *Malassezia* seen in library 3 and at lower abundance in library 5 (Table 4). As this genus of fungi is found on the skin of animals it is not surprising that it was present in these data, as all the fleas studied were collected while feeding on mammalian wildlife.

4. Discussion

We used a meta-transcriptomic approach in an initial attempt to characterize the virome of fleas (*Siphonaptera*). Accordingly, RNA-sequencing data were generated for fleas of three genera, *Pygiopsylla, Macropsylla* and *Stephanocircus*, collected across two locations along the east coast of Australia (including metropolitan Sydney) from trapped native marsupials, native bush rats, and introduced black rats. Despite their ability to transmit a number of important human pathogens such as *Yersinia pestis, Bartonella henselae, Rickettsia typhus* and tick-borne encephalitis virus, to date there has been no study of the flea virome (Azad et al., 1997; Chouikha and Hinnebusch, 2012; Rehacek, 1961). Hence, this study provides a first insight into the diversity of flea viruses as well as their association with Australian wildlife. This is of increasing importance as urban development will increase interactions among native wildlife, humans and domestic animals, facilitating cross-species virus transmission.

Despite the relatively small number of fleas sampled, we identified 18 novel RNA virus species belonging to nine families: *Picornaviridae, Sobemoviridae, Tombusviridae, Reoviridae, Narnaviridae, Phenuiviridae, Chuviridae, Totiviridae* and *Orthomyxoviridae*. No previously described virus species were identified within the data set. All 18 viruses were divergent from those viruses described previously, with Mortimer virus showing the highest level of sequence similarity to its most closely related virus at just 55% amino acid sequence similarity in the RdRp that

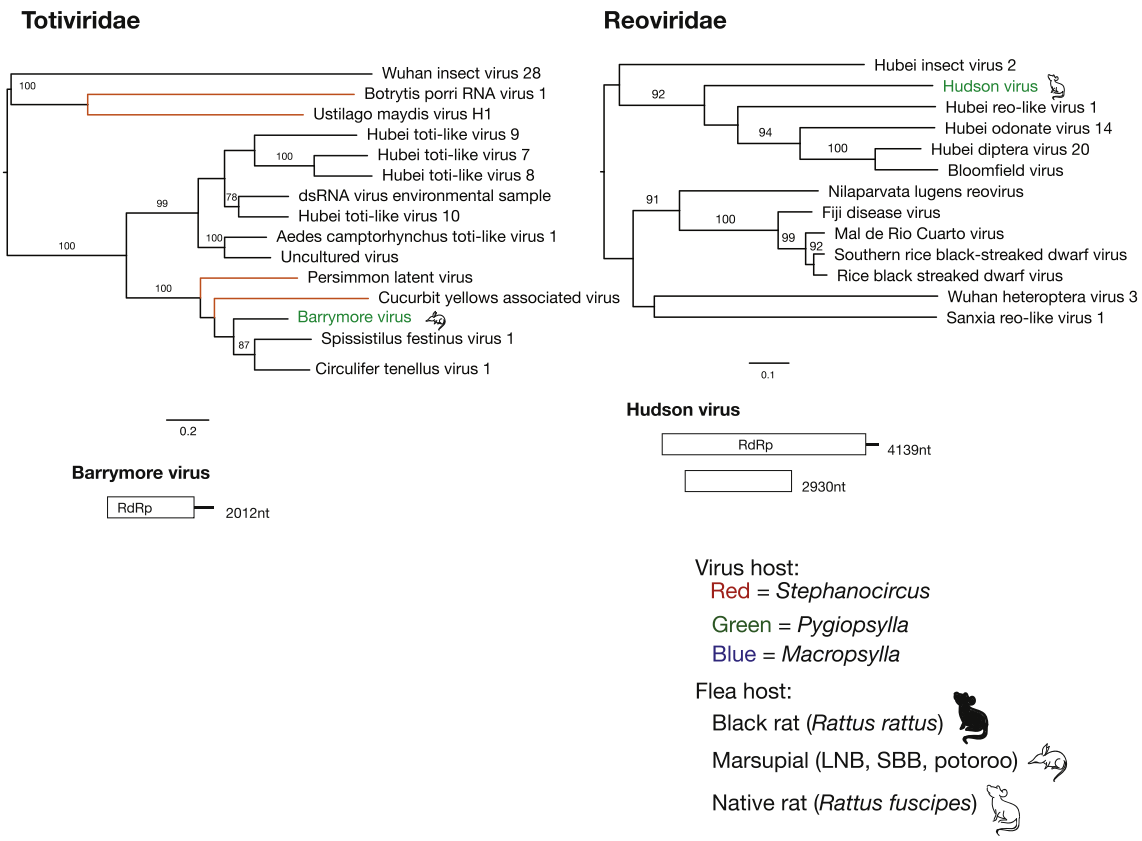

Fig. 6. Phylogenetic relationships and genomic structures of the novel dsRNA viruses. All trees are scaled to the number of amino acid substitutions per site and midpoint rooted for clarity. Bootstrap values of more than 70% are shown. Viruses shown in blue are associated with the genus *Macropsylla*, while those shown in green and red are associated with *Pygiopsylla* and *Stephanocircus* genera, respectively. Branch tips are colored according to virus hosts, with fungi and plant viruses indicated by an orange branch, vertebrate viruses with a purple branch, and invertebrate viruses with a black branch. Genome diagrams provide information on the number of genome segments identified, the length in nucleotides of each segment, the number of predicted ORFs, and any predicted conserved protein structures.

Table 3
The abundance of bacterial genera present in each library as measured by the abundance of 16S rRNA sequences. Abundance is measured using RPKM.

Genus	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6
Bacillus	57.812	84.200	50.307	–	13.930	101.468
Planococcus	5.491	–	–	–	–	–
Shigella	34.419	–	–	–	–	–
Uncultured bacterium	7.419	30.223	82.700	70.189	1.0131	–
Cardinium	–	102.345	–	–	–	–
Micrococcus	–	–	4.723	–	–	–
Streptomyces	23.430	–	2.669	–	–	–
Bartonella	–	–	–	42.772	1.261	–
Escherichia	–	–	–	40.341	–	71.353
Pseudomonas	13.027	–	–	–	2.706	–
Rhodopseudomonas	–	–	–	–	4.753	–
Serratia	–	–	–	–	48.564	–
Staphylococcus	–	–	–	–	48.967	–

is the most conserved protein among RNA viruses. Notably, all the viruses described here clustered with other invertebrate associated viruses, with a large majority of these being arthropod associated. Despite an apparent lack of association with vertebrate viruses, this does not necessarily imply that all viruses described here are strictly arthropod-specific, as a small number of viruses within particular families, such as the *Orthomyxoviridae, Picornaviridae* and *Reoviridae*, have previously been shown to infect vertebrates (Docherty and Slota, 1988; Williams et al., 2018; Yinda et al., 2016) and because the sample size was relatively small. The two most closely related reference sequences to the Phenui-like virus described here, Browner virus, are described as uncharacterized genes of the insect species *Ostrinia furnacalis* and *Pararge aegeria*. However, based on the phylogenetic relationship of the predicted amino acid sequences and the length and structure of the gene it is highly likely that these sequences are in fact derived from a virus infecting the insect used to generate the host transcriptomes.

A number of segmented viruses were identified within our study, with multiple segments identified for five of the 18 novel viruses. Identifying all the segments from an individual virus using meta-transcriptomic data can be challenging, and in some cases likely impossible when viruses are extremely divergent and/or at very low abundance within the sample. Although the number of segments in a viral genome can be predicted to some extent by the number seen in closely related viruses, this is difficult with highly divergent viruses in which segment numbers can vary greatly, even within the same family (Shi et al., 2016). Our data suggests that newly identified *Orthomyxo*-like virus

Table 4
The abundance of eukaryotic microbial organisms (protozoa and fungi) identified using CCMetagen. Abundance is measured using RPKM.

Genus	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6
Albugo	–	–	0.671	–	–	–
Aspergillus	–	1.888	–	–	–	–
Blechomonas	32.621	–	–	–	74.657	–
Cladosporium	1.146	–	15.588	–	–	–
Dictyostelium	–	–	–	0.881	–	–
Glomeromycota	1.355	–	–	–	–	–
Hepatozoon	–	–	–	–	1.724	–
Heteroconium	–	–	2.173	–	–	–
Ichthyosporea	–	–	2.879	–	–	–
Leptomonas	68.041	–	–	–	–	–
Malassezia	–	–	25.683	–	0.583	–
Oidiodendron	–	–	–	–	–	1.130
Penicillium	–	3.533	–	–	–	–
Phoma	–	–	0.706	–	–	–
Trypanosoma	–	1.428	219.022	–	–	1.149

Lestrade virus has five segments, as does its closest sampled relative, Photinus pyralis orthomyxo-like virus 2 (Fallon et al., 2018), although viruses in the *Orthomyxoviridae* family are known to have up to 8 segments (McDonald et al., 2016). Similarly, viruses in the *Reoviridae* family are known to vary in their number of segments, usually between 8 and 12, although only two segments of Hudson virus could be identified here (McDonald et al., 2016). This is likely due to the highly divergent nature of Hudson virus, exhibiting only 28% amino acid sequence similarity to the most closely related RdRp sequence, as well as its low abundance in the library.

Libraries containing fleas of the genus *Pygiopsylla* exhibited the highest diversity and abundance of virus species, including the most abundant viruses observed here (a orthomyxo-like virus and a picorna-like virus). However, library 6, comprising *Pygiopsylla* fleas collected while feeding on invasive Black rats, did not contain the abundance or diversity seen in libraries 2 and 3, also from *Pygiopsylla* but collected from native mammals. This is consistent with a study that found that Australian fleas feeding on domestic animals exhibited significantly lower microbiome diversity than those feeding on native wildlife (Lawrence et al., 2015). *Stephanocircus* fleas showed the least diversity and abundance of virus species.

As noted above, *Pygiopsylla* libraries 2 and 3 contained the two most abundant viruses as a percentage of total reads - the orthomyxo-like virus, Lestrade virus and the picorna-like Stamford virus. Lestrade virus was the most abundant within our data set as a whole, constituting 0.2% of the reads in library 2. This is not uncharacteristic of orthomyxo-like viruses, including those seen in other invertebrates (Harvey et al., 2019; Shi et al., 2017). The two least abundant viruses identified within the data set as a whole - the toti-like Barrymore virus and the narna-like Moriarty virus - were also found in libraries 2 and 3. Both virus families have been observed at low abundance in other invertebrate viromes (Shi et al., 2017). This low abundance, combined with its phylogenetic relationship to viruses found in diverse hosts, suggest that the narna-like and toti-like viruses identified here are in fact likely infecting fungal or unicellular eukaryotes associated with fleas.

A number of the viruses described here group phylogenetically with viruses identified in our previous study of Australian ticks (Harvey et al., 2019) collected during the same population studies as the fleas included here. It should be noted, however, that although both types of parasite were collected at the same time and from the same mammalian species, none of the fleas included in this study were collected from the same individual hosts as the ticks analyzed previously. Interestingly, +ssRNA viruses were more abundant in fleas while -ssRNA viruses were more common in ticks (Harvey et al., 2019). In particular, we observed a high abundance of picorna-like species in fleas, although these are commonplace in invertebrates (Shi et al., 2016). Within the *Orthomyxoviridae* phylogeny, Lestrade virus sits close to a cluster of viruses containing Old Quarry Swamp virus from *Ixodes holocyclus* ticks collected from the same region as the fleas in this study, and both were sampled from trapped marsupials. Similarly, within the *Narnaviridae*, Brunton virus grouped more closely with Yambulla virus found in *I. holocyclus* fed on southern brown bandicoots than with the other narna-like viruses found in fleas. Nadgee virus, which was also found in *I. holocyclus* ticks fed on southern brown bandicoots, falls within a sister group to Sherlock and Moriarty viruses. Notably, all those Australian tick viruses that were related to the flea viruses identified here were found in *I. holocyclus* ticks collected from bandicoots. At face value these data suggest that the mammalian hosts of these ectoparasites may in part determine the diversity of virus species observed in parasitic invertebrates, although more data is needed to confirm this. Finally, within the tick virome, three mammalian associated viruses were identified, inferred through their phylogeny and, in the case of the two flaviviruses, also their abundance. Notably, no such mammalian associated viruses were identified within our flea transcriptomes. Despite the smaller sample size, it is tempting to speculate that the difference between ticks and fleas is the smaller volume of host blood held within the latter at the time of RNA extraction. Indeed, *I. holocyclus* ticks are capable of carrying large quantities of blood comparative to their body size.

An earlier study of the microbiome of the native Australian flea species *Echidnophaga a. ambulans* found that they carried far less bacterial diversity than the common cat flea (*Ctenocephalides f. felis*) (Lawrence et al., 2015). This is consistent with the bacterial diversity seen in native flea species analyzed here, which was highest in fleas of the genus *Stephanocircus;* although, interestingly, they also harbored the least diversity and abundance of viruses. Also similar to prior work was that we identified a number of soil associated bacteria such as *Bacillus* in a variety of flea species (five of the six libraries) but not in the *Macropsylla* library. Interestingly, the flea associated bacterial genus, *Bartonella*, was identified in libraries from *Macropsylla* and *Stephanocircus* fleas. These 16S rRNA hits exhibited their closest sequence similarity to *Bartonella japonica* and *Bartonella rattaustraliani,* respectively. Interestingly, *B. japonica* is a rodent associated bacterium identified in Japanese field mice (Inoue et al., 2010), while *B. rattaustraliani* was first identified in native Australian rats (Gundi et al., 2009).

We also identified a number of eukaryotic microbial organisms in our data, with the most abundant being from the genera *Trypanosoma, Blechomonas and Leptomonas* (order *Trypanosomatida),* all of which are associated with fleas (de Avelar et al., 2011; Votýpka et al., 2013). Interestingly, the libraries in which the *Trypanosomatida* were highly abundant (libraries 1, 3 and 5) were the same libraries containing the narnaviruses and tombusviruses identified here. These viruses have previously been shown to infect trypanosomatids, with a recent study

finding that members of the *Narnaviridae* were associated with *Blechomonas spp.* (Akopyants et al., 2016; Grybchuk et al., 2018). Although a majority of the viruses that cluster with Sherlock virus, Moriarty virus and Cushing virus are listed as infecting the arthropod host in whose transcriptome they were identified, such as barnacles, crabs, shrimp, ticks and cockroaches, the increasing number of narnaviruses associated with the trypanosomatid parasites of these insects suggests that these viruses, and those identified here, may also in fact infect trypanosomatids.

In sum, by implementing a meta-transcriptomic approach we present the first virome-scale study of fleas, demonstrating that these animals are capable of carrying a wide diversity of RNA viruses and other microorganisms. Although four of the 18 novel viruses identified here clustered phylogenetically within the same groups as viruses identified in *I. holocyclus* ticks, no likely mammalian-associated viruses were identified. It therefore remains to be determined whether these parasitic vectors can transmit medically important viruses.

Conflicts of interest

The authors declare no conflicts of interests.

Acknowledgments

The project was funded by an Australian Research Council (ARC) Australian Laureate Fellowship to ECH (FL170100022).

References

Akopyants, N.S., Lye, L.F., Dobson, D.E., Lukeš, J., Beverley, S.M., 2016. A narnavirus in the trypanosomatid protest plant pathogen *Phytomonas serpens*. Genome Announc. 4, e00711–e00716.

Azad, A.F., Radulovic, S., Higgins, J.A., Noden, B.H., Troyer, J.M., 1997. Flea-borne rickettsioses: ecologic considerations. Emerg. Infect. Dis. 3, 319–327.

Bitam, I., Dittmar, K., Parola, P., Whiting, M.F., Raoult, D., 2010. Fleas and flea-borne diseases. Int. J. Infect. Dis. 14, e667–e676.

Brinkmann, A., Dincer, E., Polat, C., Hekimoglu, O., Hacioglu, S., Foldes, K., Ozkul, A., Oktem, I.M.A., Nitsche, A., Ergunay, K., 2018. A metagenomic survey identifies Tamdy orthonairovirus as well as divergent phlebo-, rhabdo-, chu- and flavi-like viruses in Anatolia, Turkey. Ticks Tick-Borne Dis. 9, 1173–1183.

Brinkmann, A., Nitsche, A., Kohl, C., 2016. Viral metagenomics on blood-feeding arthropods as a tool for human disease surveillance. Int. J. Mol. Sci. 17, E1743.

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

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Chouikha, I., Hinnebusch, B.J., 2012. *Yersinia*-flea interactions and the evolution of the arthropod-borne transmission route of plague. Curr. Opin. Microbiol. 15, 239–246.

Clausen, P.T., Aarestrup, F.M., Lund, O., 2018. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinf. 19, 307.

de Alevar, D.M., Melo, M.N., Linardi, P.M., 2011. Morphology and growth characteristics of cultures *Leptomonas ctenocephali* from *Ctenocephalides felis felis* (Siphonaptera: Pulicidae) of dogs in Brazil. Vet. Parasitol. 180, 394–398.

Docherty, D.E., Slota, P.G., 1988. Use of muscovy duck embryo fibroblasts for the isolation of viruses from wild birds. J. Tissue Cult. Methods 11, 165–170.

Dunnet, G., Nardon, D., 1974. A monograph of Australian fleas (Siphonaptera). Aust. J. Zool.; Suppl. Ser. 22, 1–273.

Eads, D.A., Hoogland, J.L., 2017. Precipitation, climate change, and parasitism of prairie dogs by fleas that transmit Plague. J. Parasitol. 103, 309–319.

Fallon, T.R., Lower, S.E., Chang, C.H., Bessho-Uehara, M., Martin, G.J., Bewick, A.J., Behringer, M., Debat, H.j., Wong, I.,, Day, J.C., Suvorov, A., Silva, C.J., Stanger-Hall, K.F., Hall, D.W., Schmitz, R.J., Nelson, D.R., Lewis, S.M., Shigenobu, S., Bybee, S.M., Larracuente, A.M., Oba, Y., Weng, J.K., 2018. Firefly genomes illustrate parallel origins of bioluminescence in beetles. eLife 7, e36495.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Fridman, N., Regev, A., 2011. Full-length transcriptome assembly for RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Grybchuk, D., Kostygov, A.Y., Macedo, D.H., Votýpka, J., Lukeš, J., Yurchenko, V., 2018. RNA viruses in *Blechomonas* (Trypanosomatidae) and evolution of *Leishmaniavirus*. mBio 9 e01932-18.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Gundi, V.A., Taylor, C., Raoult, D., La Scola, B., 2009. *Bartonella rattaustraliani* sp. *nov.*, *Bartonella queenslandensis* sp. nov. and *Bartonella coopersplainsensis* sp. nov., identified in Australian rats. Int. J. Syst. Evol. Microbiol. 59, 2956–2961.

Harvey, E., Rose, K., Eden, J.S., Lo, N., Abeyasuriya, T., Shi, M., Doggett, S.L., Holmes, E.C., 2019. Extensive diversity of RNA viruses in Australian ticks. J. Virol. 93 e01358-18.

Inoue, K., Kabeya, H., Shiratori, H., Ueda, K., Kosoy, M.Y., Chomel, B.B., Boulouis, H.J., Maruyama, S., 2010. *Bartonella japonica* sp. *nov.* and *Bartonella silvatica* sp. *nov.*, isolated from *Apodemus* mice. Int. J. Syst. Evol. Microbiol. 60, 759–763.

Kaewmongkol, G., Kaewmongkol, S., Burmej, H., Bennett, M.D., Fleming, P.A., Adams, P.J., Wayne, A.F., Ryan, U., Irwin, P.J., Fenwick, S.G., 2011. Diversity of *Bartonella* species detected in arthropod vectors from animals in Australia. Comp. Immunol. Microbiol. Infect. Dis. 34, 411–417.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kerr, P.J., Liu, J., Cattadori, I., Ghedin, E., Read, A.F., Holmes, E.C., 2015. Myxoma virus and the leporipoxviruses: an evolutionary paradigm. Viruses 7, 1029–1061.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Lawrence, A.L., Hii, S.F., Chong, R., Webb, C.E., Traub, R., Brown, G., Slapeta, J., 2015. Evaluation of the bacterial microbiome of two flea species using different DNA-isolation techniques provides insights into flea host ecology. FEMS Microbiol. Ecol. 91, fiv134.

Marcelino, V.R., Clausen, P.T., Buchman, J., Wille, M., Iredell, J.R., Meyer, W., Lund, O., Sorrell, T., Holmes, E.C., 2019. CCMetagen: comprehensive and accurate identification of eukaryotes and prokaryotes in metagenomics data. bioRxiv. https://doi.org/10.1101/641332.

Mencke, N., Vobis, M., Mehlhorn, H., J, D.H., Rehagen, M., Mangold-Gehring, S., Truyen, U., 2009. Transmission of feline calicivirus via the cat flea (*Ctenocephalides felis*). Parasitol. Res. 105, 185–189.

McDonald, S., Nelson, M.I., Turner, P.E., Patton, J.T., 2016. Reassortment in segmented RNA viruses: mechanisms and outcomes. Nat. Rev. Microbiol. 14, 448–460.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Oshima, Y., Fujii, M., Shiogama, K., Miyamoto, K., Fujita, H., Sato, S., Maruyama, S., Mahara, F., Tsutsumi, Y., 2016. *Bartonella henselae* infection caused by cat flea bite. Pathol. Int. 66, 177–179.

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

Raoult, D., Mouffok, N., Bitam, I., Piarroux, R., Drancourt, M., 2013. Plague: history and contemporary analysis. J. Infect. 66, 18–26.

Rehacek, J., 1961. Transmission of tick-borne encephalitis virus by fleas. J. Hyg. Epidemiol. Microbiol. Immunol. (Prague) 5, 282–285.

Schloderer, D., Owen, H., Clark, P., Stenos, J., Fenwick, S.G., 2006. *Rickettsia felis* in fleas, Western Australia. Emerg. Infect. Dis. 12, 841–843.

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

Shi, M., Neville, P., Nicholson, J., Eden, J.S., Imrie, A., Holmes, E.C., 2017. High-resolution metatranscriptomics reveals the ecological dynamics of mosquito-associated RNA viruses in Western Australia. J. Virol. 91 e00680-17.

Tokarz, R., Williams, S.H., Sameroff, S., Sanchez Leon, M., Jain, K., Lipkin, W.I., 2014. Virome analysis of *Amblyomma americanum, Dermacentor variabilis*, and *Ixodes scapularis* ticks reveals novel highly divergent vertebrate and invertebrate viruses. J. Virol. 88, 11480–11492.

Trebbien, R., Chriel, M., Struve, T., Hjulsager, C.K., Larsen, G., Larsen, L.E., 2014. Wildlife reservoirs of canine distemper virus resulted in a major outbreak in Danish farmed mink (*Neovison vison*). PLoS One 9, e85598.

Vobis, M., D'Haese, J., Mehlhorn, H., Mencke, N., 2003. Evidence of horizontal transmission of feline leukemia virus by the cat flea (*Ctenocephalides felis*). Parasitol. Res. 91, 467–470.

Van der Mescht, L., Matthee, S., 2017. Host range and distribution of small mammal fleas in South Africa, with a focus on species of medical and veterinary importance. Med. Vet. Entomol. 31, 402–413.

Votýpka, J., Suková, E., Kraeva, N., Ishemgulova, A., Duží, I., Lukeš, J., Yurchenko, V., 2013. Diversity of trypanosomatids (Kinetoplastea: Trypanosomatidae) parasitizing fleas (Insecta: Siphonaptera) and description of a new genus *Blechomonas* gen. n. Protist 164, 763–781.

Williams, S.H., Che, X., Garcia, J.A., Klena, J.D., Lee, B., Muller, D., Ulrich, W., Corrigan, R.M., Nichol, S., Jain, K., Lipkin, W.I., 2018. Viral diversity of house mice in New York city. mBio 9 e01354-17.

Yinda, C.K., Zeller, M., Conceicao-Neto, N., Maes, P., Deboutte, W., Beller, L., Heylen, E., Ghogomu, S.M., Van Ranst, M., Matthijnssens, J., 2016. Novel highly divergent reassortant bat rotaviruses in Cameroon, without evidence of zoonosis. Sci. Rep. 6, 34209.

Zhang, Y.Z., Shi, M., Holmes, E.C., 2018. Using metagenomics to characterize an expanding virosphere. Cell 172, 1168–1172.