DR. MICHELLE WILLE (Orcid ID : 0000-0002-5629-0196)

Article type : Original Article

Virus-virus interactions and host ecology are associated with RNA virome structure in wild birds

Short Title: RNA virome in wild birds

Michelle Wille[1,*], John-Sebastian Eden[2,3], Mang Shi[2], Marcel Klaassen[4], Aeron C. Hurt[1], Edward C. Holmes[2,*]

[1]WHO Collaborating Centre for Reference and Research on Influenza, at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia.
[2]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, Australia.
[3]The Westmead Institute for Medical Research, Centre for Virus Research, Westmead, Australia.
[4]Centre for Integrative Ecology, Deakin University, Geelong, Australia.

*Corresponding author: Michelle Wille michelle.wille@influenzacentre.org. Edward C. Holmes edward.holmes@sydney.edu.au

Abstract

Little is known about the factors that shape the ecology of RNA viruses in nature. Wild birds are an important case in point, as other than influenza A virus, avian samples are rarely tested for viruses, especially in the absence of overt disease. Using bulk RNA-sequencing ('meta-transcriptomics') we revealed the viral diversity present in Australian wild birds through the lens of the ecological factors that may determine virome structure and abundance. A meta-transcriptomic analysis of four Anseriformes (waterfowl) and Charadriiformes (shorebird) species sampled in temperate and arid Australia revealed the presence of 27 RNA virus genomes, 18 of which represent newly described species. The viruses identified included a previously described gammacoronavirus and influenza A viruses. Additionally, we identified novel virus species from the families *Astroviridae*, *Caliciviridae*, *Reoviridae*, *Rhabdoviridae*, *Picobirnaviridae*, and *Picornaviridae*. We noted differences in virome structure that reflected underlying differences in location and influenza A infection status. Red-necked avocets (*Recurvirostra novaehollandiae*) from Australia's arid interior possessed the greatest viral diversity and abundance, markedly higher than individuals sampled in temperate Australia. In Ruddy Turnstones (*Arenaria interpres*) and dabbling ducks (*Anas* spp.) viral abundance and diversity was higher and more similar in hosts that were positive for influenza A infection compared to those that were negative for this virus, despite samples being collected on the same day and from the same location. This study highlights the extent and diversity of RNA viruses in wild birds, and lays the foundation for understanding the factors that determine virome structure in wild populations.

Keywords: virus evolution; ecology; host - pathogen interactions; influenza A virus; virome; wild birds

Introduction

Wild birds are ubiquitous, found on every continent, and a massive biomass of these animals move across the globe on annual cycles of migration creating a truly interconnected planet (Bauer & Hoye 2014). In addition to natural environments, birds can be found in our cities, using waste water treatment plants, landfills, and our drinking water reservoirs. Beyond wild birds, it is estimated that 1 out of every 7-14 birds on earth are raised for human consumption (*i.e.* chickens) (Barrowclough *et al.* 2016; Food and Agriculture Organization of the United

Nations 2012), which may act as important amplifiers of potentially zoonotic avian viruses, such as influenza A virus (IAV) (Gao *et al.* 2013; Wan 2012; Yoon *et al.* 2015). Despite our important relationship with birds, we have only a limited understanding of the diversity of avian viruses. Indeed, most studies of avian viruses have focused on those that cause mass mortality in wild birds (e.g. Wellfleet bay virus (Allison *et al.* 2014; Ballard *et al.* 2017)), result in substantial economic losses in food production birds (e.g. avian avulavirus type 1 (Alexander 2000; Leighton & Heckert 2007; Tolf *et al.* 2013)), or are zoonotic (e.g. IAV (Gao *et al.* 2013; Wan 2012; Yoon *et al.* 2015)).

Avian viruses have a rich and complex ecology (van Dijk *et al.* 2018), with patterns of prevalence affected by seasonality (Latorre-Margalef *et al.* 2014), host species (Munster *et al.* 2007), host age (van Dijk *et al.* 2014), latitude (Lisovski *et al.* 2017), and urbanization (Wille *et al.* 2017). However, although most studies of virus ecology and evolution have implicitly assumed a "one-host, one-virus" model of host-pathogen interactions, such as the Mallard (*Anas platyrhynchos*) – IAV model (van Dijk *et al.* 2014), both hosts and their viruses exist in communities, and it is likely that these communities are the drivers of viral ecology. For example, despite intensive focus on the Mallard – IAV system (Latorre-Marglalef *et al*. 2014, van Dijk *et al*. 2014), IAV is in reality a multi-host virus detected in over 100 species of wild birds, with different avian species playing different roles in virus ecology and evolution (Olsen *et al.* 2006). For example, gulls are reservoirs for evolutionary distinct IAV subtypes (Wille *et al.* 2011), and rare subtypes may be maintained in members of the Anseriiformes and Charadriiformes that are infrequently sampled (Wille *et al.* 2018). In turn, numerous viruses have been detected in wild bird populations, and Mallards maybe co-infected with at least three different RNA viruses simultaneously in the absence of overt signs of disease (Wille *et al.* 2015; Wille *et al.* 2017). These other avian RNA viruses – avian coronavirus and avian avulavirus type 1 – have seasonal prevalence patterns that generally mirror that of IAV, and it is therefore possible that they may also share similar host taxonomic or geographic differences in viral community structure (Wille *et al.* 2017). Virus co-infection may also be an important driver of viral prevalence, as virus-virus interactions may enhance or interfere with infection (Diaz-Munoz 2017; Elena & Sanjuan 2005; Henle 1950; Jolly & Narayan 1989). As a case in point, a higher prevalence of avian coronavirus was found in a bird population experiencing IAV infection, suggesting that the latter might play a role in structuring avian viromes in general (Wille *et al.* 2015).

Although one in 10 bird species is found in Australia, we know little of the accompanying viral diversity and abundance in these animals, nor of the large-scale ecological factors that determine virome composition. We used a recently developed unbiased meta-transcriptomic pipeline based on bulk RNA-sequencing (Shi *et al.* 2018b) to reveal the viromes of four Australian avian species, and to evaluate how the structure of entire viral communities is impacted by a variety of ecological correlates. In particular, we assessed the role of host taxonomy, location, and co-infection with IAV on virome structure. This study illustrates the extent of RNA viral diversity in wild birds, and the importance of the expansion of traditional host-pathogen systems beyond simple one-host, one-virus systems to disentangle ecological processes in viral presence and abundance.

Materials and Methods

Ethics statement

This research was conducted under approval of the Deakin University Animal Ethics Committee (permit numbers A113-2010 and B37-2013). Banding was performed under Australian Bird Banding Scheme permit (banding authority numbers 2915 and 2703). Research permits were approved by the Department of Environment, Land, Water and Planning Victoria (permit numbers 10006663 and 10005726), Department of Environment, Water and Natural Resources South Australia (research permit numbers M25919-1,2,3,4,5) and the Department of Primary Industries, Parks, Water and Environment Tasmania (permit number FA 13032).

Sample selection

Samples were collected as part of a long-term IAV surveillance study (Ferenczi 2016; Ferenczi *et al.* 2016). Birds were captured using baited funnel walk-in traps, cannon nets or mist nets as described previously (Ferenczi 2016). Importantly, none of the birds in this study exhibited any signs of disease. Samples were collected from (i) three temperate locations in Australia – the Western Treatment Plant near Melbourne (37°59′11.62″S, 144°39′38.66″E), Yallock Creek (38°13'51.6"S 145°28'43.9"E), King Island (39°55′52″S 143°51′02″E), and (ii) an interior arid location – Innamincka Regional Reserve (27°32′28″S 140°35′47″E). Species selected for the study included both those known to be important in IAV ecology (*Anas* ducks and Ruddy Turnstone) and those in which IAV has not been described (Australian Shelduck and Red-necked Avocet) (Table 1).

Cloacal samples (from 2012, 2013) or a combination or oropharangeal and cloacal samples (from 2014) were collected using a sterile-tipped swab and were placed in viral transport media (VTM, Brain-heart infusion broth containing $2x10^6$ U/l penicillin, 0.2 mg/ml streptomycin, 0.5 mg/ml gentamicin, 500 U/ml amphotericin B, Sigma). All samples were assayed for IAV as described previously (Ferenczi *et al.* 2016).

RNA library construction and sequencing

RNA was extracted using the MagMax *mir*Vana™ Total RNA isolation Kit (ThermoFisher Scientific) on the KingFisher™ Flex Purification System (ThermoFisher Scientific). All extracted samples were assessed for RNA quality using the TapeStation 2200 and High Sensitivity RNA reagents (Aligent Genomics, Integrated Sciences), and 10 samples with the highest concentration were pooled (based on species, location, and IAV infection status) using equal concentrations and subsequently concentrated using the RNeasy MinElute Cleanup Kit (Qiagen). Libraries were constructed using the TruSeq total RNA library preparation protocol (Illumina) and host rRNA was removed using the Ribo-Zero-Gold kit (Illumina). Paired end sequencing (100bp) of the RNA library was performed on the HiSeq2500 platform. All library preparation and sequencing was carried out at the Australian Genome Research Facility (AGRF, Melbourne).

RNA virus discovery

Sequence reads were demultiplexed and trimmed with Trimmomatic followed by *de novo* assembly using Trinity (Grabherr *et al.* 2011). No filtering of host/bacterial reads was performed before assembly. All assembled contigs were compared to the entire non-redundant nucleotide (nt) and protein (nr) database using blastn and diamond blast (Buchfink *et al.* 2015), respectively, setting an e-value threshold of $1x10^{-10}$ to remove potential false positives. Abundance estimates for all contigs were determined using the RSEM algorithm implemented in Trinity. All contigs that returned blast hits with paired abundance estimates were filtered to remove plants and invertebrate reads that likely correspond to the host diet, as well as fungal, bacterial and host sequences. Blast results were used to initially classify viruses to their appropriate family and genus level, and the virus list was further filtered to remove viruses with invertebrate (Shi *et al.* 2016), plant or bacterial host associations using the Virus-Host database (http://www.genome.jp/virushostdb/).

To compare relative viral abundance across libraries, three host reference genes were mined from the data using a custom blast database. As not all host reference genes are stably expressed, we utilized three genes that are stably expressed in the Mallard (*Anas playrhynchos*) lower gastrointestinal tract (Chapman *et al.* 2016): Ribosomal Protein L4 (RPL4), Ribosomal Protein S13 (RPS13) and NADH dehydrogenase 1 alpha subcomplex (NDUFA) from both Mallard (taxid: 8839), Chicken (*Gallus gallus*) (taxid: 9031) and Zebra Finch (*Taeniopygia guttata*) (taxid: 59729).

Virus genome annotation and phylogenetic analysis

Contigs greater than 1000bp in length were inspected using Geneious R10 (Biomatters, New Zealand), and open reading frames corresponding to predicted genome architectures based on the closest reference genomes were interrogated using the conserved domain database (CDD, https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) with an e-value threshold of $1x10^{-5}$. Reads were subsequently mapped back to viral contigs to identify mis-assembly using the Geneious mapping function. Viruses with full length genomes, or incomplete genomes but that possess the full-length RNA-dependant RNA polymerase (RdRp) gene, were used for phylogenetic analysis. Briefly, amino acid sequences of the polyprotein or gene encoding for the RdRp were aligned using MAFFT (Katoh & Standley 2013), and gaps and ambiguously aligned regions were stripped using trimAL (Capella-Gutierrez *et al.* 2009). Final alignment lengths are presented in Table S2. The most appropriate amino acid substitution model was then determined for each data set, and maximum likelihood trees were estimated using PhyML 3.0 (Guindon *et al.* 2010) with 1000 bootstrap replicates using the ATGC server (http://www.atgc-montpellier.fr/phyml/execution.php). For IAV and gammacoronavirus, phylogenies were also estimated using the nucleotide sequences of complete or partial reference genome sequences to better place viruses in context of currently described avian viral diversity. Similarly, the best-fit model of nucleotide substitution was selected, and maximum likelihood trees were estimated using PhyML 3.0 with 1000 bootstrap replications. Novel viral species were identified as those that had <90% RdRp protein identity, or <80% genome identity to previously described viruses.

Diversity and abundance across libraries

Relative virus abundance was estimated as the proportion of the total viral reads in each library (excluding rRNA). All ecological measures were calculated using the data set comprising "higher" vertebrate-associated viruses (i.e. those associated with birds and mammals), albeit with all retroviruses and retrovirus-like elements removed (hereafter, avian virus data set). Analyses were performed using *R* version 3.4.0 integrated into RStudio v 1.0.143, and plotted using *ggplot2*. Specifically, both the observed richness and Shannon effective [alpha diversity] were calculated for each library at the family and genus levels using the *Rhea* script sets (Lagkouvardos *et al.* 2017). Beta diversity was calculated using the Bray Curtis dissimilarity matrix using the *vegan* package (Oksanen *et al.* 2007) at both family and genus levels, and presented as a network using the *igraph* (Csardi & Nepisz 2006) and *ggnet* (https://github.com/briatte/ggnet) packages. Non-metric multidimensional scaling (NMDS) ordination was additionally calculated based on bray-curtis dissimilarity, and Adonis tests (PERMANOVA) applied using the *phyloseq* package (McMurdie & Holmes 2013).

Results

RNA-Seq as a means to identify avian viruses

We characterized the total transcriptome of eight avian pools, representing two tribes in the order Anseriiformes (waterfowl) and two families in the order Charadriiformes (shorebirds). These pools were designed to answer specific questions on the determinants of virome structure including the impact of bird taxonomy (within and between Anseriiformes and Charadriiformes), location (temperate versus arid sampling sites) and effect of IAV infection (Table 1). Each library comprised swab samples collected from 10 individuals at the same time point and location, therefore increasing the chances of finding viruses at lower prevalence.

RNA sequencing of rRNA depleted libraries resulted in a median of 44,345,130 (range 39,267,372 – 47,650,666) reads per pool, which were assembled into a median of 175,559 contigs (range 135,254 – 357,869). An assessment of the host reference gene RPS13, a proxy for sequencing depth of libraries, revealed similar abundances (0.000102-0.000342% of reads), suggesting similar host sequencing depth across the libraries (Fig S1). These eight libraries had marked differences in the abundance of avian viral reads; Ruddy Turnstones

(*Arenaria interpres*) and *Anas* ducks that were IAV positive (0.21% and 0.1% of reads), and Red-necked Avocets (*Recurvirostra novaehollandiae*) from the interior (0.26% of reads) had relatively high abundances of avian viruses, while lower abundance levels were observed in Ruddy Turnstones and *Anas* ducks that were IAV negative (0.00006% and 0.00051% of reads, respectively) (Fig 1).

Blast analysis and virus characterization revealed the genomes of 27 RNA viruses, of which 18 were newly described species – that is, of sufficient phylogenetic distinction to represent new virus species) – but that belonged to existing families and were most closely related to other avian viruses (Table S1). The viruses identified comprised double-stranded RNA viruses (*Reoviridae*, Rotavirus; *Picobirnaviridae*), positive-sense single-stranded RNA viruses (*Caliciviridae*; *Astroviridae*; *Picornaviridae*, genus *Avihepatovirus*, *Megrivirus* and Unassigned genera; *Coronaviridae*, genus *Gammacoronavirus*) and negative-sense single-stranded RNA viruses (*Rhabdoviridae*; genus *Tupavirus*; *Orthomyxoviridae*, genus Influenza A virus). Members of the family *Paramyxoviridae*, known to circulate in wild birds (Ramey *et al.* 2013; Wille *et al.* 2015), were notably absent. No DNA viruses (*i.e.* the RNA transcripts of DNA viruses), were detected, potentially because cloacal and oropharangeal samples are a richer source of shed viruses rather than those actively replicating within cells of the gastrointestinal tract. An array of retroviruses, or retrovirus-like elements were also detected, but due to the challenge in differentiating between endogenous and exogenous retroviruses they will not be discussed here (Fig 1).

Substantial undescribed diversity of RNA viruses in wild birds

An average of 80% of virus species in each library were novel (range 50-100%) and in three libraries all viruses were novel (Table S1), illustrating the large undiscovered viral diversity in wild birds. Numerous new viral species were identified from viral families that are not frequently screened for in wild birds, including rhabdoviruses, caliciviruses, picornaviruses and rotaviruses. In the case of some viral families, complete viruses were only found in a single species, such as the picobirnaviruses detected in Australian Shelducks (*Tadorna tadornoides*) (Fig 2A). Other viruses, such as the caliciviruses, were highly abundant across all avian hosts, and full genomes were found in all avian species included in this study (Fig 4, Fig S9).

Across all the RNA-dependant RNA polymerase (RdRp) phylogenies, the viral species from wild birds generated in this study were in the most part similar to previously described avian viruses, often forming an apparent "avian" clade within each group of viruses and suggestive of a relatively long-term association with birds (Figs 2-6). Although there were some exceptions, such as the Shelduck picornaviruses which were most closely related to those viruses sampled in swine viruses, it is possible that this simply reflects poor sampling.

Wild birds have previously been described as hosts for coronaviruses, astroviruses, and IAVs, and the viruses identified in this study belonged to "wild bird" clades to which sequences from poultry fell as outgroups (S3-S8 Figs). Specifically, four different IAV HA-NA subtypes were found in this study; H12N5 and H9N3 viruses from *Anas* ducks, and H6N8 and H10N8 from Ruddy Turnstones (S3-S5 Figs). In addition to H12N5 and H9N3 in *Anas* ducks, one short contig from the HA of H10 influenza virus was also identified. These three subtypes, all identified in *Anas* ducks, had different abundances in this library: H12N5 comprised 77% of all avian viral reads (0.081% of all reads in the library), compared to H9N3 that comprised 0.5% of avian viral reads in the library (0.0005% of all reads), and H10 that represented only 0.0068% of avian reads. This is in comparison to the H6N8 and H10N8 viruses identified in the Ruddy Turnstones that had similar abundances (25-35% of avian viral reads, 0.05-0.07% of all reads in the library).

Broadly, the IAVs from Australian birds described here were most similar to viruses sampled from Eurasian wild birds, which is expected given that Australian migratory birds use the East-Asian-Australian flyway. However, while the H12 virus was more similar to viruses from Eurasia, this virus was phylogenetically distinct from currently circulating viruses, suggesting the presence of a potential "Australia-specific" clade. Additionally, the Ruddy Turnstone H10 sequence fell into the North American clade rather than the Eurasian clade, in contrast to the N8 segment which fell into a Eurasian clade. Such a phylogenetic pattern is indicative of the intercontinental reassortment of these Ruddy Turnstone viruses. The gammacoronavirus in this study, identified in Red-necked Avocets, was related to circulating wild bird gammacoronaviruses from waterfowl from both Eurasia and the United States (S7-S8 Figs).

Novel dsRNA viruses. Two complete picobirnavirus genomes were found in wild birds samples (Fig 2A). These viruses, all from Australian Shelducks, clustered together on the phylogenetic tree, although a more divergent partial virus was also found. Australian Shelducks are particularly rich in picobirnavirus diversity, with 21% of all avian viral reads in the library derived from picobirnaviruses (0.002% of total reads), although this virus family was also found at low abundance in *Anas* ducks from the interior. This potentially suggests a preference for the Anseriiformes, although this will need to be confirmed with more data (Fig 2A, Fig S9). In addition, rotaviruses were found in almost all libraries, and in all host groups (Fig S9) with three different subtypes revealed (Fig 2B). Specifically, we found previously described rotavirus D and G viruses in apparently healthy ducks and avocets, respectively, even though they are known to cause enteritis in poultry. Unlike other phylogenies in which wild bird viruses fell in clades that are distinct from poultry-associated viruses, wild bird and poultry rotaviruses were phylogenetically similar and hence clustered on the tree, although sample size was limited. Ruddy Turnstones, however, carried a highly abundant (53% of avian viral reads, 0.11% of total reads in the library) and highly divergent rotavirus, characterized by a long branch, that fell as a sister-species to rotavirus G (Fig 2B).

Novel ssRNA viruses. Two novel avastroviruses were identified in Red-necked Avocets, both falling as outgroups to Group 2 viruses, the archetype of which is Avian Nephritis virus (Fig 3, Fig S6). These viruses share 60% and 40% pairwise amino acid identity to Avian Nephritis virus, respectively, suggesting that there is a large undescribed diversity of wild bird avastroviruses. Calicivirus reads were identified in all libraries, with the exception of *Anas* ducks from the interior of Australia. Furthermore, full genomes of five caliciviruses were identified in four libraries at high abundance (Australian Shelducks 0.005% of total reads, Ruddy Turnstone IAV positive 0.006% of total reads, *Anas* duck IAV positive 0.002% of total reads, Avocet Interior 0.03% of total reads), and all these viruses belonged to the same clade as currently described poultry viruses within an unassigned genus (Fig 4). Two novel rhabdoviruses from Anseriiformes found in this study fell as a divergent group within the genus Tupaviruses, within which Durham virus is the only avian virus previously described (Fig 5). Specifically, Shelduck rhabdovirus and Duck rhabdovirus fell as relatively distantly related sister species, as illustrated by long branch lengths on the phylogeny, and thereby potentially represent a novel clade of wild bird viruses.

Our virome sampling also revealed a great diversity of picornaviruses (*Picornaviridae*), almost all of which were found in Red-necked Avocets (Fig 6). Megriviruses were remarkably abundant in avocets, comprising 30% and 65% of all avian viral reads (0.004% and 0.169% of total reads) from birds sampled in two locations, representing two locations. Furthermore, megriviruses represent the only virus (other than IAV) found in more than one library (Figs 6, 7). The library generated from avocets in the interior also contained (i) Avocet picornavirus B-like A, that formed a sister group to pigeon picornavirus B, (ii) Avocet picornavirus B-like B, a sister group to a clade containing both pigeon picornavirus B and Avocet picornavirus B-like A, and (iii) Avocet picornavirus, a highly divergent virus that likely represents a novel genus. Finally, in *Anas* ducks, we observed a divergent sister group to Duck hepatitis A 1 and 3 (Wild Duck avihepatovirus-like) (Fig 6).

Factors affecting the structure and abundance of avian viromes

One of the most important results of our study was that IAV status andlocation, but not host taxonomy, was associated with differences in viral abundance and diversity. Because of the potential impact of host phylogeny, we expected that virome structure would be similar within the Anseriiformes and Charadriiformes but differ between these orders. However, across all the libraries and controlling for location and IAV status, libraries from members of the Anseriiformes were no more similar to each other than they were to those from the Charadriiformes. This lack of taxonomic distinction was apparent whether the analysis was performed at the level of viral species, genus, or family (Figs 7, S10-S12, viral family, R^2 = 0.142, p = 0.353, viral genus R^2 = 0.153, p = 0.251), although the sample size was relatively small. To better control for other variables in our sampling scheme, we compared the libraries from avocets and dabbling ducks as these were all IAV negative and sampled from the same locations; in this case the relationship remains statistically insignificant (viral family, R^2=0.25, p=0.4; viral genus R^2=0.28, p=0.3), although we lose statistical power due to the small sample size. Furthermore, the three libraries from *Anas* ducks had a different abundance and viral composition, and these three libraries represent differences in IAV infection status and location (Fig 7, Figs S10-S12). No viral species were shared within host species or family, with the exception of IAV and a megrivirus found in both avocet libraries (Fig 7). Finally, there were no differences in viral family distribution at the level of host species and order; all viral families were found in both Anseriiformes and Charadriiformes, with the exception of picobirnaviruses which only occurred in the Anseriiformes (Fig S12).

It might also been expected that birds in temperate locationss would have a higher viral abundance and diversity than birds from the arid interior of Australia given greater prevalence of IAV in temperate latitudes (Lisovski *et al.* 2017). However, the library from Red-necked Avocets from the arid interior had a higher viral abundance and diversity than individuals from temperate Australia, and also had the highest viral abundance across all libraries (0.26% total reads) (Figs 1,7-8, Figs S10-S11). There was also a clear virome difference between *Anas* ducks across locations: a higher viral abundance and species diversity was found in the ducks from the interior compared to the temperate ducks that were negative for IAV. However, temperate ducks that were IAV positive had a higher viral diversity and abundance compared to ducks sampled from the interior (Fig 8A-B). Overall, incorporating all libraries and controlling for IAV infection status and host species, location did not predict higher similarity between the libraries, as libraries from the same location were no more similar to each other than those from different locations (viral family, $R^2 = 0.1$, $p = 0.554$, viral genus $R^2=0.093$, $p=0.8$) (Figs 7,8, Figs S10-S12). When comparing only dabbling ducks that were IAV negative and avocets from arid and temperate locations, the relationship remained statistically insignificant (viral family $R^2=0.199$, $p=0.9$, viral genus $R^2=0.2$, $p=1$); however, there was limited statistical power due to small sample size.

Finally, we expected that, in accord with previous studies (Wille *et al.* 2015; Wille *et al.* 2017), libraries containing IAV would have higher viral diversity compared to those that were negative for IAV. To address this, samples from IAV positive and negative birds were selected from the same location during the same sampling expedition to remove any potential bias. Libraries from both Ruddy Turnstone and *Anas* ducks that were positive for IAV indeed had a higher viral abundance (Fig 8A) [0.21%, 0.1% compared to 0.000061%, 0.0005% viral reads] and virus diversity (Figs 7, 8C, Figs S10-S11), at the family, genus and species levels (Fig 7, 8C, Figs S10-S11). This trend remained when abundance or diversity attributable to IAV was removed from the analysis. Furthermore, the two libraries containing IAV were statistically significantly more similar to each other in abundance and composition compared to all other libraries sequenced (viral family including IAV reads, $R^2 = 0.24$, $p = 0.008$, viral genus including IAV reads, $R^2=0.24$, $p=0.014$, viral family excluding IAV reads, $R^2=0.17$, $p=0.04$, viral genus excluding IAV reads $R^2=0.24$, $p=0.017$) (Figs 7,8, Figs S11-S12). Only including the Turnstone and *Anas* duck libraries negatively affects statistical power due to small sample size.

Discussion

Despite the ubiquitous nature and economic importance of birds, we have a poor understanding of the natural viral diversity in this major animal phylum. To this end, we employed an unbiased metagenomics approach to reveal avian viromes, comprising 27 novel and previously described viral species, in a framework of ecological hypothesis testing.

Given the long-term association between hosts and viruses, it was not unexpected that the viruses revealed in this study were most closely related to other avian viruses (Shi *et al.* 2018a), especially virulent poultry viruses that have been an important focus in virus characterization efforts (Boros *et al.* 2016; Day *et al.* 2010). Based on previous studies we anticipated finding low pathogenic forms of coronaviruses, astroviruses, and avulaviruses (Wille *et al.* 2015; Wille *et al.* 2017). While we did detect most of these viruses, the absence of avian avulavirus was surprising. Avian avulavirus type 1 is present in wild birds in Australia (Hoque *et al.* 2012a; Hoque *et al.* 2012b; Hore *et al.* 1973; Mackenzie *et al.* 1984; Peroulis & O'Riley 2004) and globally (Alexander 2000; Austin & Webster 1993; Hanson *et al.* 2005; Ramey *et al.* 2013; Tolf *et al.* 2013; Wille *et al.* 2015), and has been detected in co-infection studies (Wille *et al.* 2015; Wille *et al.* 2017). We also genomically described 18 new viral species which belonged to previously identified avian clades (genera) predominantly comprised of poultry viruses. For example, we identified five caliciviruses from wild birds that belonged to a previously described avian clade in the *Caliciviridae* comprised of chicken, turkey and goose caliciviruses (Liao *et al.* 2014; Wolf *et al.* 2012; Wolf *et al.* 2011).

While virus species and genotypes that were sister groups to poultry viruses were revealed, it is important to note all samples in this study were collected from birds with no clinical signs of disease. This raises two important issues. First, with the sequencing of more wild birds, those clades formally dominated by poultry will likely expand to include many viral species and genotypes from wild birds. This will be central to a better understanding the movement of avian viruses between wild bird reservoirs and poultry populations and hence of disease emergence in general. Indeed, poultry production has rapidly expanded in the last century (Kaleta & Rulke 2008), to the extent that ~70% of avian biomass on the planet are now poultry (Bar-On *et al.* 2018), creating a relatively new, but large niche for viruses. Furthermore, unlike wild birds, in viruses adapted to poultry such as Marek's Disease virus (a

double-stand DNA virus), there has likely been selection for high transmissibility and high virulence (Rozins & Day 2017). As such, sequencing wild bird viruses is imperative in understanding the evolutionary processes involved.

Second, these data raise the issue of how wild birds are able to tolerate such high levels of virus diversity and abundance seemingly in the absence of overt disease (Medzhitov *et al.* 2012; Råberg 2014). In particular, cloacal swabs of Red-necked avocets from the interior had a high viral abundance (0.26% of reads were from avian viruses) and these 10 birds shed 13 viral genera and eight viral species for which full genomes were revealed. Ruddy Turnstone and *Anas* ducks that were infected with IAV similarly shed avian viruses at high abundance (0.21% and 0.1%, respectively), albeit with lower viral diversity. Although this must impose some physiological effect on the host, there continues to be conflicting data on the physiological effect of IAV infection in isolation (Kuiken 2013), let alone the viral abundance described in this study. A large viral diversity in healthy, individual wild birds (Fawaz *et al.* 2016; Wille *et al.* 2015) and poultry (Day *et al.* 2010; Day *et al.* 2015; Lima *et al.* 2017) has been previously described, in contrast to chickens in which viral diversity was described in diseased animals (e.g. diarrhoea; (Boros *et al.* 2016)). The leading hypothesis reflects a long history of host-pathogen co-evolution: chickens are a relatively new host niche for IAV, and following spill-over from wild birds highly pathogenic phenotypes evolve (for example, H7 (Seekings *et al.* 2018), sometimes resulting in catastrophic mortality. This is in contrast to wild birds that have likely been co-evolving with IAV over long time periods, with natural selection perhaps favouring lower virulence; as a consequence, the highly pathogenic IAV found in wild birds are most likely due to spill over from poultry (Barber *et al.* 2008; Little *et al.* 2010; van Dijk *et al.* 2015). Indeed, a muted inflammatory response translating to immunological tolerance to viral infections, may allow some hosts, such as bats, to harbour a variety of viruses (Brook & Dobson 2015; Xie *et al.* 2018). Similarly, Pekin Ducks (*Anas platyrhynchos domesticus*) and wild strain Mallard ducks appear to have a controlled innate immune response against both low pathogenic IAV (Helin *et al.* 2018) and highly pathogenic IAV (Saito *et al.* 2018), with upregulation of the innate immune system occurring on day 1 post infection and no evidence of hypercytokinemia, or "cytokine storms". Given the high (>0.1% of reads) viral abundance in some libraries, it is possible that some of the viruses described here do not cause disease in the absence of some other physiological or environmental stressor, although this is clearly an issue that needs to be explored in more detail.

Viral co-infection is likely to be the rule rather than the exception, and is shaped by both host ecology and virus-virus interactions (Diaz-Munoz 2017). To date, much of what we know about viruses in wild birds is derived from many years of research in IAVs (Ferenczi *et al.* 2016; Latorre-Margalef *et al.* 2014; Munster *et al.* 2007; Olsen *et al.* 2006) and those RNA viruses that have similar patterns of host preference and seasonality (Wille *et al.* 2015; Wille *et al.* 2017). Members of the Anseriiformes and Charadriiformes have proven to be excellent model species for this study, and we detected viruses previously described in these hosts (Chu *et al.* 2011; Muradrasoli *et al.* 2010; Wille *et al.* 2017; Wille *et al.* 2016) as well as a suite of novel viruses. One of the key observations of our study is that avian taxonomy did not drive virome structure; that is, there was no specific clustering of libraries in the NMDS plots by avian order (Anseriiformes versus Charadriiformes), and the three *Anas* duck libraries were different, although this analysis had limited statistical power. Given that IAV is prevalent in both these avian orders, we suggest that host ecology may play a more important role than host taxonomy in shaping virome diversity. For example, it is possible that waterbirds share viral families, genera and species as shallow water bodies facilitating virus transmission between individuals, as with IAV (Hoye *et al.* 2012; van Dijk *et al.* 2018). In support of this there was a difference between birds sampled (Red-necked Avocets in particular) in lakes of the arid interior as compared to temperate coastlines.

The ecological factors assessed here are not mutually exclusive, as shown by the complex relationship between the three *Anas* duck libraries which had different structures given different conditions (location and IAV status). Samples from Ruddy Turnstones were collected from the same beaches on the same sampling trip, yet the 10 birds positive and negative for IAV had different viromes (total abundance, species abundances, and viral diversity). Furthermore, the viromes of *Anas* ducks and Ruddy Turnstones that were positive for IAV were more similar to each other than to all other libraries. Anecdotally, while we were able to successfully culture H6N8 and H10N8 virus as part of an ongoing IAV surveillance project, the viruses from the *Anas* ducks were not successfully isolated, demonstrating the power of the meta-transcriptomic approach used here. These trends, however, may be biased due sample pooling, such that the patterns may be due to only a few individuals in the pool. Ultimately, therefore, the validity of the patterns observed here need to be re-assessed on the basis of individual transcriptomes, although such work will obviously be both costly and time-consuming. In addition, it was previously shown (Wille *et al.* 2015) that co-infection with IAV was important in modulating the prevalence of other RNA viruses.

In particular, IAV modulates the interferon response of the host, changing the antiviral state (Garcia-Sastre 2001, 2011; Hale *et al.* 2008), and this may promote co-infection or prevent viral clearance of certain viruses. We have a surprisingly poor understanding of virus-virus interactions, although viruses do have mechanisms to mediate infection by other viruses. For example, viruses may have synergistic (enhancing) or antagonistic (inhibiting) interactions, and this may occur within and/or across viral species (Diaz-Munoz 2017; Elena & Sanjuan 2005; Henle 1950; Jolly & Narayan 1989). Regardless, virus-virus interactions are important drivers of co-infection, and may be further affected by virus and host ecology (Diaz-Munoz 2017).

In sum, we have expanded our understanding of the diversity of avian viruses and laid the foundation for future hypothesis testing of the factors associated with virome structure in wild birds using high throughput metatranscriptomics. This study focused on avian orders known to be central in the ecology of IAV, but also a number of other RNA viruses. Although this study is of a relatively limited scale from an ecological perspective, we have successfully described viral diversity in samples collected from different sites, times, and avian species and found evidence for differences across these factors.. Finally, we demonstrate several potential applications of viral community analyses, and anticipate a rapid expansion of viral ecology to move beyond the one, host – one virus, system, and to consider both viruses and hosts as complex ecological communities.

Acknowledgements

The sampling of birds in Australia importantly relied on the assistance of volunteers within the Victorian Wader Study Group, members of the Centre for Integrative Ecology at Deakin University, notably Simeon Lisovski, Marta Ferenczi, and David Roshier, and the logistic support from Melbourne Water, Innamincka Station, Gidgealpa Station and Innamincka Regional Reserve.

The sampling was supported by NIH/NIAID (HHSN266200700010C), ARC discovery grants (DP 130101935 and DP160102146) and an ARC Australian Laureate Fellowship to ECH (FL170100022). The WHO Collaborating Centre for Reference and Research on Influenza is funded by the Australian Commonwealth Government.

References

Alexander DJ (2000) Newcastle disease and other avian paramyxoviruses. *Rev Sci Tech OIE* **19**, 443-462.

Allison AB, Ballard JR, Tesh RB, *et al.* (2014) Cyclic avian mass mortality in the northeastern United States is associated with a novel orthomyxovirus. *J Virol*, doi: 10.1128/JVI.02019-02014.

Austin FJ, Webster RG (1993) Evidence of ortho- and paramyxoviruses in fauna from Antarctica. *J Wildl Dis* **29**, 568-571.

Ballard JR, Mickley R, Gibbs SEJ, *et al.* (2017) Prevalence and distribution of Wellfleet Bay virus exposure in the Common Eider (*Somateria mollissima*). *J Wildl Dis* **53**, 81-90.

Bar-On YM, Phillips R, Milo R (2018) The biomass distribution on Earth. *Proc Natl Acad Sci USA*, 201711842. doi:201711810.201711073/pnas.1711842115

Barber MRW, Aldridge JR, Webster RG, Magor KE (2008) Association of RIG-I with innate immunity of ducks to influenza. *Proc Natl Acad Sci USA* **107**, 5913-5918.

Barrowclough GF, Cracraft J, Klicka J, Zink RM (2016) How many kinds of birds are there and why does it matter? *PLoS ONE* **11**, e0166307. doi: 0166310.0161371/journal.pone.0166307.

Bauer S, Hoye BJ (2014) Migratory animals couple biodiversity and ecosystem functioning worldwide. *Science* **344**, 54. doi: 10.1126/Science.1242552.

Boros A, Pankovics P, Adonyi A, *et al.* (2016) A diarrheic chicken simultaneously co-infected with multiple picornaviruses: Complete genome analysis of avian picornaviruses representing up to six genera. *Virology* **489**, 63-74.

Brook CE, Dobson AP (2015) Bats as 'special' reservoirs for emerging zoonotic pathogens. *Trends Microbiol* **23**, 172-180.

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nature Methods* **12**, 59-60.

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

Chapman JR, Helin AS, Wille M, *et al.* (2016) A panel of stably expressed reference genes for real-time qPCR gene expression studies of Mallards (*Anas platyrhynchos*). *PLoS ONE* **11**, e0149454. doi: 0149410.0141371/journal.pone.0149454.

Chu DK, Leung CY, Gilbert M, *et al.* (2011) Avian coronavirus in wild aquatic birds. *J Virol* **85**, 12815-12820.

Csardi G, Nepisz T (2006) The igraph software package for complex network research. *Complex Syst*, 1695. http://igraph.org.

Day JM, Ballard LL, Duke MV, Scheffler BE, Zsak L (2010) Metagenomic analysis of the turkey gut RNA virus community. *Virol J* **7**, 313. doi: 310.1186/1743-1422X-1187-1313.

Day JM, Oakley BB, Seal BS, Zsak L (2015) Comparative analysis of the intestinal bacterial and RNA viral communities from sentinel birds placed on selected broiler chicken farms. *PLoS ONE* **10**, e0117210.

Diaz-Munoz SL (2017) Viral coinfection is shaped by host ecology and virus-virus interactions across diverse microbial taxa and environments. *Virus Evol* **3**.

Elena SF, Sanjuan R (2005) RNA viruses as complex adaptive systems. *Bio Systems* **81**, 31-41.

Fawaz M, Vijayakumar P, Mishra A, *et al.* (2016) Duck gut viral metagenome analysis captures snapshot of viral diversity. *Gut Pathog* **8**, 30.

Ferenczi M (2016) *Avian influenza virus dynamics in Australian wild birds. PhD Thesis*,

Deakin University.
Ferenczi M, Beckmann C, Warner S, *et al.* (2016) Avian influenza infection dynamics under variable climatic conditions, viral prevalence is rainfall driven in waterfowl from temperate, south-east Australia. *Vet Res* **47**, 23.
Food and Agriculture Organization of the United Nations (2012) *FAOSTAT Database* FAO, Rome, Italy.
Gao R, Cao B, Hu Y, *et al.* (2013) Human infection with a novel avian-origin influenza A (H7N9) virus. *The New England journal of medicine* **368**, 1888-1897.
Garcia-Sastre A (2001) Inhibition of interferon-mediated antiviral responses by influenza A viruses and other negative-strand RNA viruses. *Virology* **279**, 375-384.
Garcia-Sastre A (2011) Induction and evasion of type I interferon responses by influenza viruses. *Virus Res* **162**, 12-18.
Grabherr MG, Haas BJ, Yassour M, *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotech* **29**, 644-652.
Guindon S, Dufayard JF, Lefort V, *et al.* (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321.
Hale BG, Randall RE, Ortin J, Jackson D (2008) The multifunctional NS1 protein of influenza A viruses. *The Journal of general virology* **89**, 2359-2376.
Hanson BA, Swayne DE, Senne DA, *et al.* (2005) Avian influenza viruses and paramyxoviruses in wintering and resident ducks in Texas. *J Wildl Dis* **41**, 624-628.
Helin AS, Wille M, Atterby C, *et al.* (2018) A rapid and transient innate immune response to avian influenza infection in mallards. *Mol Immunol* **95**, 64-72.
Henle W (1950) Interference phenomena between animal viruses - a review. *J Immunol* **64**, 203-236.
Hoque MA, Burgess GW, Greenhil AR, Hedlefs R, Skerratt LF (2012a) Causes of morbidity and mortality of wild aquatic birds at billabong sanctuary, Townsville, North Queensland, Australia. *Av Dis* **56**, 249-256.
Hoque MA, Burgess GW, Karo-Karo D, Cheam AL, Skerratt LF (2012b) Monitoring of wild birds for Newcastle disease virus in north Queensland, Australia. *Prev Vet Med* **103**, 49-62.
Hore DE, Campbell J, Turner AJ (1973) A serological survey for viral antibodies in wild ducks. *Australian veterinary journal* **49**, 238-239.
Hoye BJ, Fouchier RAM, Klaassen M (2012) Host behaviour and physiology underpin individual variation in avian influenza virus infection in migratory Bewick's Swans. *Proc Royal Soc B* **279**, 529-534.
Jolly PE, Narayan O (1989) Evidence for interference, coinfections, and intertypic virus enhancement of infection by ovine-caprine lentiviruses. *J Virol* **63**, 4682-4688.
Kaleta EF, Rulke CPA (2008) The beginning and spread of fowl plague (H7 High Pathogenicity Avian Influenza) across Europe and Asia (1978-1955). In: *Avian Influenza* (ed. Swayne DE), pp. 145-189. Blackwell Publishing, U.K. .
Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780.
Kuiken T (2013) Is low pathogenic avian influenza virus virulent for wild waterbirds? *Proc Royal Soc B* **280**, 20130990. doi: 20130910.20131098/rspb.20132013.20130990.
Lagkouvardos I, Fischer S, Kumar N, Clavel T (2017) Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *Peerj* **5**, e2836. doi: 2810.7717/peerj.2836.
Latorre-Margalef N, Tolf C, Grosbois V, *et al.* (2014) Long-term variation in influenza A virus prevalence and subtype diversity in a migratory Mallards in Northern Europe.

Proc Royal Soc B **281**, doi: 10.1098/rspb.2014.0098
Leighton FA, Heckert RA (2007) Newcastle Disease and Related Avian Paramyxoviruses. In: *Infectious Diseases of Wild Birds* (eds. Thomas NJ, Hunter BD, Atkinson CT), pp. 3-16. Blackwell Publishing, Ames, U.S.A.
Liao QF, Wang XY, Wang D, Zhang DB (2014) Complete genome sequence of a novel calicivirus from a goose. *Arch Virol* **159**, 2529-2531.
Lima DA, Cibulski SP, Finkler F, *et al.* (2017) Faecal virome of healthy chickens reveals a large diversity of the eukaryote viral community, including novel circular ssDNA viruses. *Journal of General Virology* **98**, 690-703.
Lisovski S, Hoye BJ, Klaassen M (2017) Geographic variation in seasonality and its influence on the dynamics of an infectious disease. *Oikos* **126**, 931-936.
Little TJ, Shuker DM, Colegrave N, Day T, Graham AL (2010) The coevolution of virulence: tolerance in perspective. *PLoS Pathog* **6**, e1001006. doi: 1001010.1001371/journal.ppat.1001006.
Mackenzie JS, Edwards EC, Holmes RM, Hinshaw VS (1984) Isolation of otho- viruses and paramyxoviruses from wild birds in Western Australia and the characterization of novel influenza A viruses. *Aust J Exp Biol Med Sci* **62**, 89-99.
McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217. doi: 61210.61371/journal.pone.0061217.
Medzhitov R, Schneider DS, Soares MP (2012) Disease tolerance as a defense strategy. *Science* **335**, 936-941.
Munster VJ, Baas C, Lexmond P, *et al.* (2007) Spatial, temporal, and species variation in prevalence of influenza A viruses in wild migratory birds. *PLoS Pathog* **3**, e61. doi: 10.1371/journal.ppat.0030061.
Muradrasoli S, Balint A, Wahlgren J, *et al.* (2010) Prevalence and phylogeny of coronaviruses in wild birds from the Bering Strait area (Beringia). *PLoS ONE* **5**, e13640. doi: 13610.11371/journal.pone.0013640.
Oksanen J, Kindt R, Legendre P, *et al.* (2007) The vegan package. *Community Ecology Package* **10**, 631-637.
Olsen B, Munster VJ, Wallensten A, *et al.* (2006) Global patterns of influenza A virus in wild birds. *Science* **312**, 384-388.
Peroulis I, O'Riley K (2004) Detection of avian paramyxoviruses and influenza viruses amongst wild bird populations in Victoria. *Australian veterinary journal* **82**, 79-82.
Råberg L (2014) How to live with the enemy: understanding tolerance to parasites. *PLoS Biol* **12**, e1001989. doi: 1001910.1001371/journal.pbio.1001989.
Ramey AM, Reeves AB, Ogawa H, *et al.* (2013) Genetic diversity and mutation of avian paramyxovirus serotype 1 (Newcastle disease virus) in wild birds and evidence for intercontinental spread. *Arch Virol* **158**, 2495-2503.
Rozins C, Day T (2017) The industrialization of farming may be driving virulence evolution. *Evol Appl* **10**, 189-198.
Saito LB, Diaz-Satizabal L, Evseev D, *et al.* (2018) IFN and cytokine responses in ducks to genetically similar H5N1 influenza A viruses of varying pathogenicity. *The Journal of general virology* **99**, 464-474. doi 410.1099/jgv.1090.001015.
Seekings AH, Slomka MJ, Russell C, *et al.* (2018) Direct evidence of H7N7 avian influenza virus mutation from low to high virulence on a single poultry premises during an outbreak in free range chickens in the UK, 2008. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **64**, 13-31.
Shi M, Lin XD, Chen X, *et al.* (2018a) The evolutionary history of vertebrate RNA viruses.

Nature **556**, 197-202.
Shi M, Lin XD, Tian JH, *et al.* (2016) Redefining the invertebrate RNA virosphere. *Nature* **540**, 539-543.
Shi M, Zhang YZ, Holmes EC (2018b) Meta-transcriptomics and the evolutionary biology of RNA viruses. *Virus Res* **243**, 83-90.
Tolf C, Wille M, Haidar A-K, *et al.* (2013) Prevalence of avian paramyxovirus type 1 in Mallards during autumn migration in the western Baltic Sea region. *Virol J* **10**, 285. doi: 210.1186/1743-1422X-1110-1285.
van Dijk JGB, Fouchier RAM, Klaassen M, Matson KD (2015) Minor differences in body condition and immune status between avian influenza virus-infected and noninfected mallards: a sign of coevolution? *Ecology and Evolution* **5**, 436-449.
van Dijk JGB, Hoye BJ, Verhagen JH, *et al.* (2014) Juveniles and migrants as drivers for seasonal epizootics of avian influenza virus. *J Anim Ecol* **83**, 266-275.
van Dijk JGB, Verhagen JH, Wille M, Waldenström J (2018) Host and virus ecology as determinants of influenza A virus transmission in wild birds. *Curr Opin Virol* **28**, 26-36.
Walker PJ, Firth C, Widen SG, *et al.* (2015) Evolution of genome size and complexity in the Rhabdoviridae. *PLoS Pathog* **11**, e1004664. doi: 1004610.1001371/journal.ppat.1004664.
Wan XF (2012) Lessons from emergence of A/goose/Guangdong/1996-like H5N1 highly pathogenic avian influenza viruses and recent influenza surveillance efforts in southern China. *Zoonoses Public Health* **59 Suppl 2**, 32-42.
Wille M, Avril A, Tolf C, *et al.* (2015) Temporal dynamics, diversity, and interplay in three components of the viriodiversity of a Mallard population: Influenza A virus, avian paramyxovirus and avian coronavirus. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **29**, 129-137.
Wille M, Latorre-Margalef N, Tolf C, *et al.* (2018) Where do all the subtypes go? Temporal dynamics of H8–H12 influenza A viruses in waterfowl. *Virus Evolution* **4**, doi: 10.1093/ve/vey1025.
Wille M, Lindqvist K, Muradrasoli S, Olsen B, Jarhult JD (2017) Urbanization and the dynamics of RNA viruses in Mallards (*Anas platyrhynchos*). *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **51**, 89-97.
Wille M, Muradrasoli S, Nilsson A, Jarhult JD (2016) High prevalence and putative lineage maintenance of avian coronaviruses in Scandinavian waterfowl. *PLoS ONE* **11**, e0150198. doi:0150110.0151371/journal.pone.0150198.
Wille M, Robertson GJ, Whitney H, *et al.* (2011) Extensive geographic mosaicism in avian influenza viruses from gulls in the northern hemisphere. *PLoS ONE* **6**, e20664. doi: 20610.21371/journal.pone.0020664.
Wolf S, Reetz J, Hoffmann K, *et al.* (2012) Discovery and genetic characterization of novel caliciviruses in German and Dutch poultry. *Arch Virol* **157**, 1499-1507.
Wolf S, Reetz J, Otto P (2011) Genetic characterization of a novel calicivirus from a chicken. *Arch Virol* **156**, 1143-1150.
Xie J, Li Y, Shen X, *et al.* (2018) Dampened STING-dependent interferon activation in bats. *Cell Host Microbe* **23**, 297-301. doi: 210.1016/j.chom.2018.1001.1006.
Yoon H, Moon OK, Jeong W, *et al.* (2015) H5N8 highly pathogenic avian influenza in the Republic of Korea: epidemiology during the first wave, from January through July 2014. *Osong Public Health Res Perspect* **6**, 106-111.

Data Accessibility

All sequence reads have been deposited in the Short Read Archive BioProject PRJNA472212. Viral genomes or complete RdRp have been deposited in GenBank, Acession MH453800 - MH453880.

Author Contributions

MW, MK, ACH, ECH designed research; MW performed research; MW, JSE, MS analysed the data; MK, ACH, ECH contributed reagents and analytical tools; MW wrote the manuscript, with input from all co-authors. All authors gave final approval for publication.

Tables and Figures

Table 1. Eight libraries were sequenced reflecting different avian taxonomy, location within Australia, and influenza A infection status.

Taxonomy	**Host Species**	**Location**	**Sampling Location**	**Influenza A status**	**Sample type**	**Year**
Anseriiformes, Anatidae, Tadorninae	Australian Shelduck (*Tadorna tadornoides*)	Temperate	Western Treatment Plant, Victoria	Negative	Cloacal	2012
Anseriiformes, Anatidae, Anatinae	Dabbling duck (*Anas* sp.)	Temperate	Western Treatment Plant, Victoria	Positive	Cloacal	2013
	Dabbling duck	Temperate	Western Treatment Plant, Victoria	Negative	Cloacal	2013
	Dabbling duck	Interior	Innamincka Regional Reserve, South Australia	Negative	Cloacal	2013
Charadriiformes, Recurvirostridae	Red-necked Avocet (*Recurvirostra novaehollandiae*)	Temperate	Yallock Creek, Victoria	Negative	Cloacal	2013
	Red-necked Avocet	Interior	Innamincka Regional Reserve, South Australia	Negative	Cloacal	2013
Charadriiformes, Scolopacidae	Ruddy Turnstone (*Arenaria interpres*)	Temperate	King Island, Tasmania	Negative	Combined oropharangeal/ cloacal	2014
	Ruddy Turnstone	Temperate	King Island, Tasmania	Positive	Combined oropharangeal/ cloacal	2014

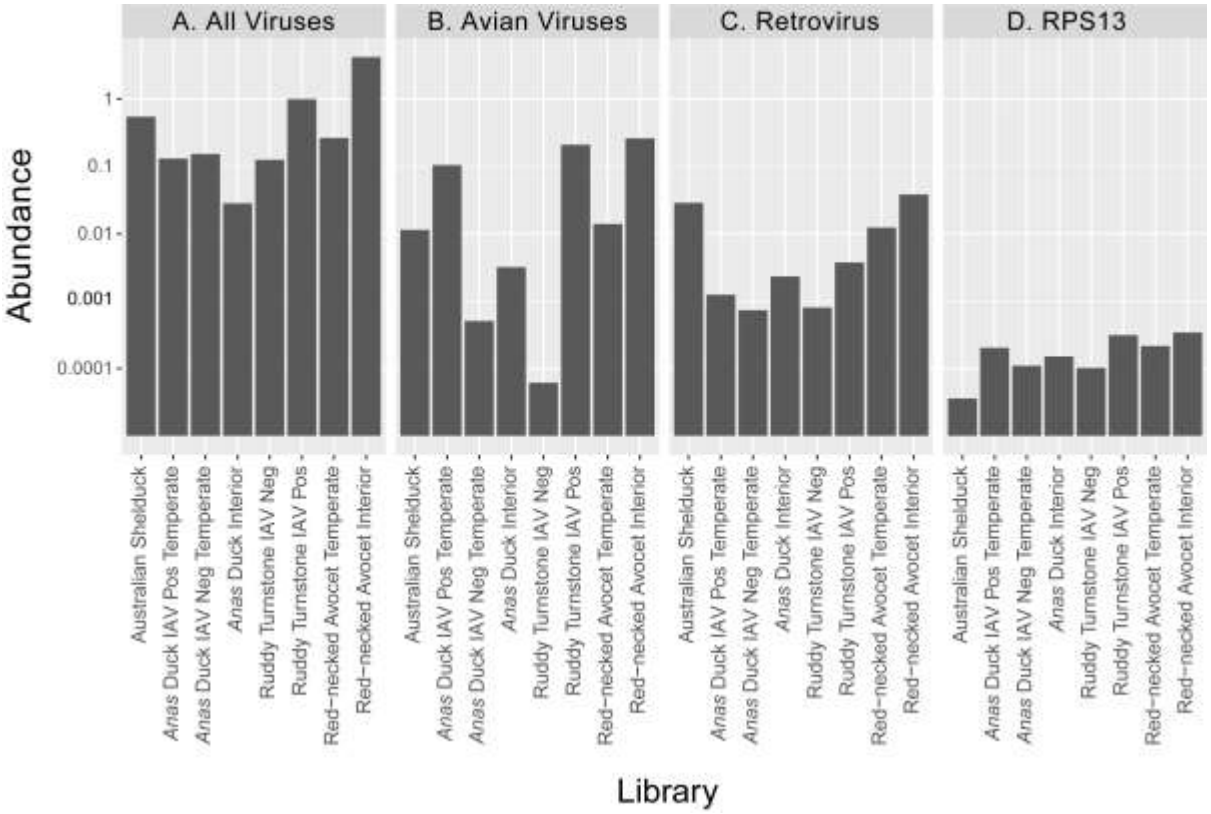

Fig 1. Abundance of viruses in each library. (A) Abundance of all exogenous viruses including those from avian, invertebrate, lower vertebrate, plant, fungi or bacterial hosts. (B) Abundance of all viruses that are associated with birds. (C) Abundance of retroviruses or retrovirus-like elements that have avian or mammalian signature. (D). Host reference gene RPS13.

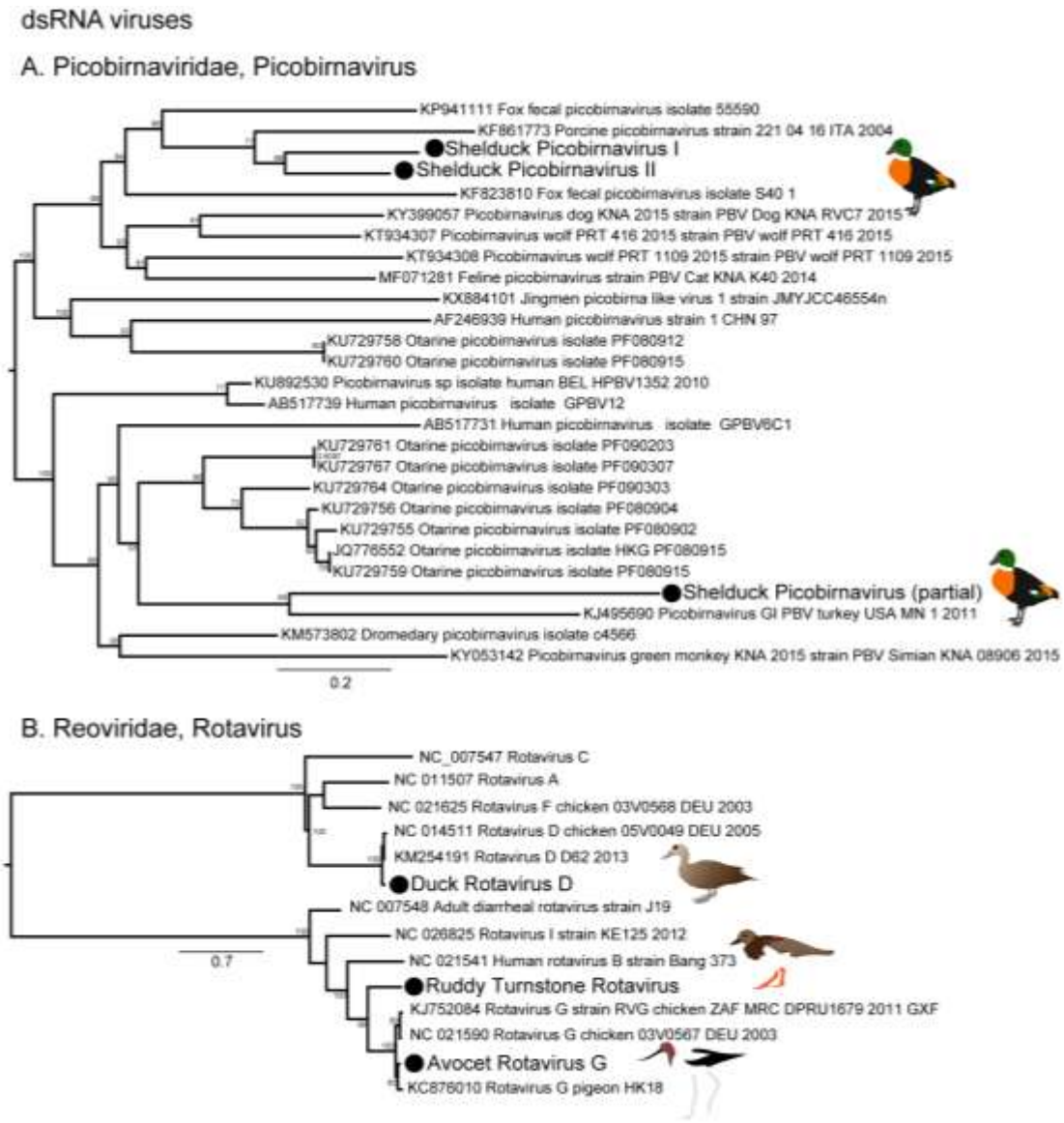

Fig 2. Phylogenies of double-stranded RNA viruses. These trees show (a) segment 2 (RdRp) of picobirnaviruses, and (b) the VP1 segment (RdRp) of the rotaviruses described in this study. All phylogenetic trees were midpoint rooted for clarity only. The scale bar indicates the number of amino acid substitutions per site. Bootstrap values >70% are shown for key nodes. Viruses described in this study are marked with a filled circle.

Fig 3. Partial RpRp phylogeny of members of the genus Avastrovirus. The tree is rooted between the avian and mammalian astroviruses. The scale bar indicates the number of nucleotide substitutions per site. Bootstrap values >70% are shown for key nodes. Viruses described in this study are marked with a filled circle. The phylogeny of the full length polyprotein is presented in Fig S6.

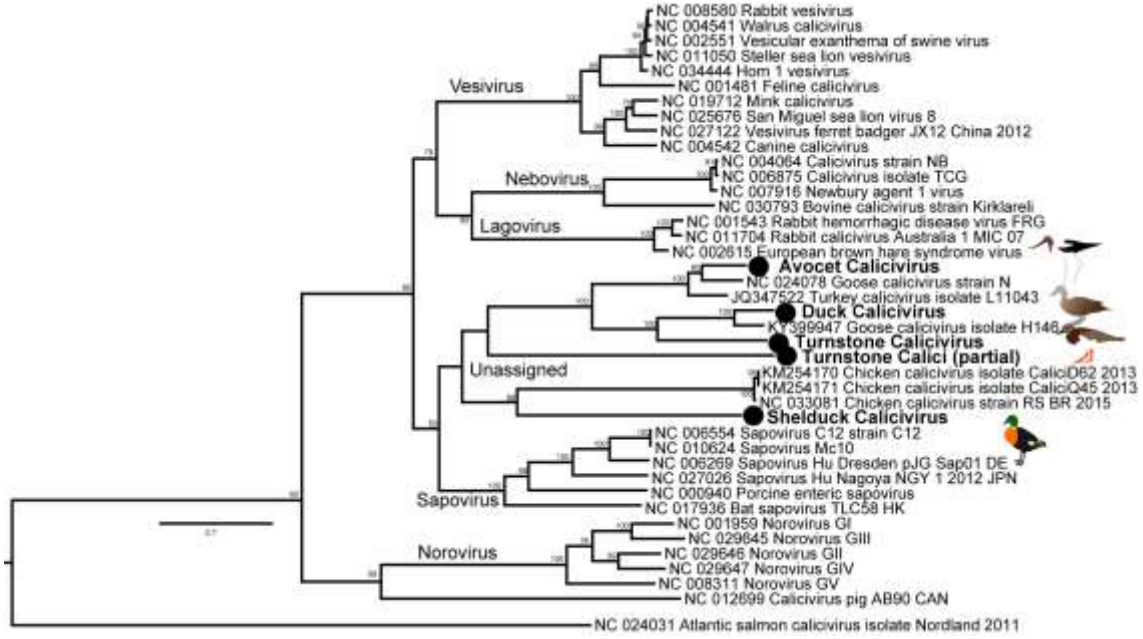

Fig 4. Phylogenetic tree of the polyprotein, containing the RdRp, of members of the *Caliciviridae*. The most divergent calicivirus, Atlantic Salmon calicivirus, was used as an outgroup to root the tree. The scale bar indicates the number of amino acid substitutions per site. Bootstrap values >70% are shown for key nodes. Viruses described in this study are marked with a filled circle.

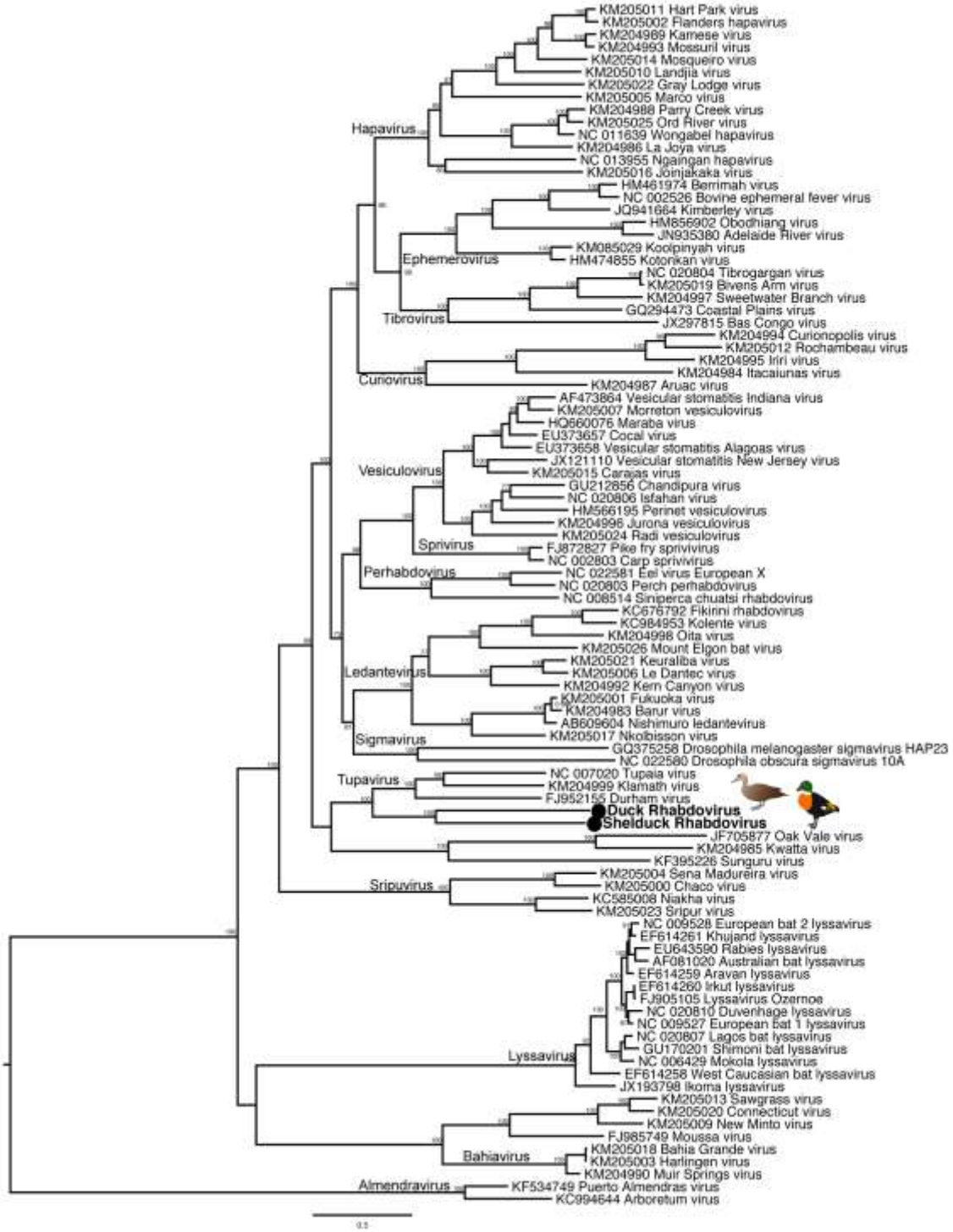

Fig 5. Phylogeny of the L gene (RdRp) of members of the *Rhabdoviridae*. Almendraviruses were set as the outgroup and representative viruses for each genus (as per (Walker *et al.* 2015)) were also included in the analysis. The scale bar indicates the number of amino acid substitutions per site. Bootstrap values >70% are shown for key nodes. Viruses described in this study are marked with a filled circle.

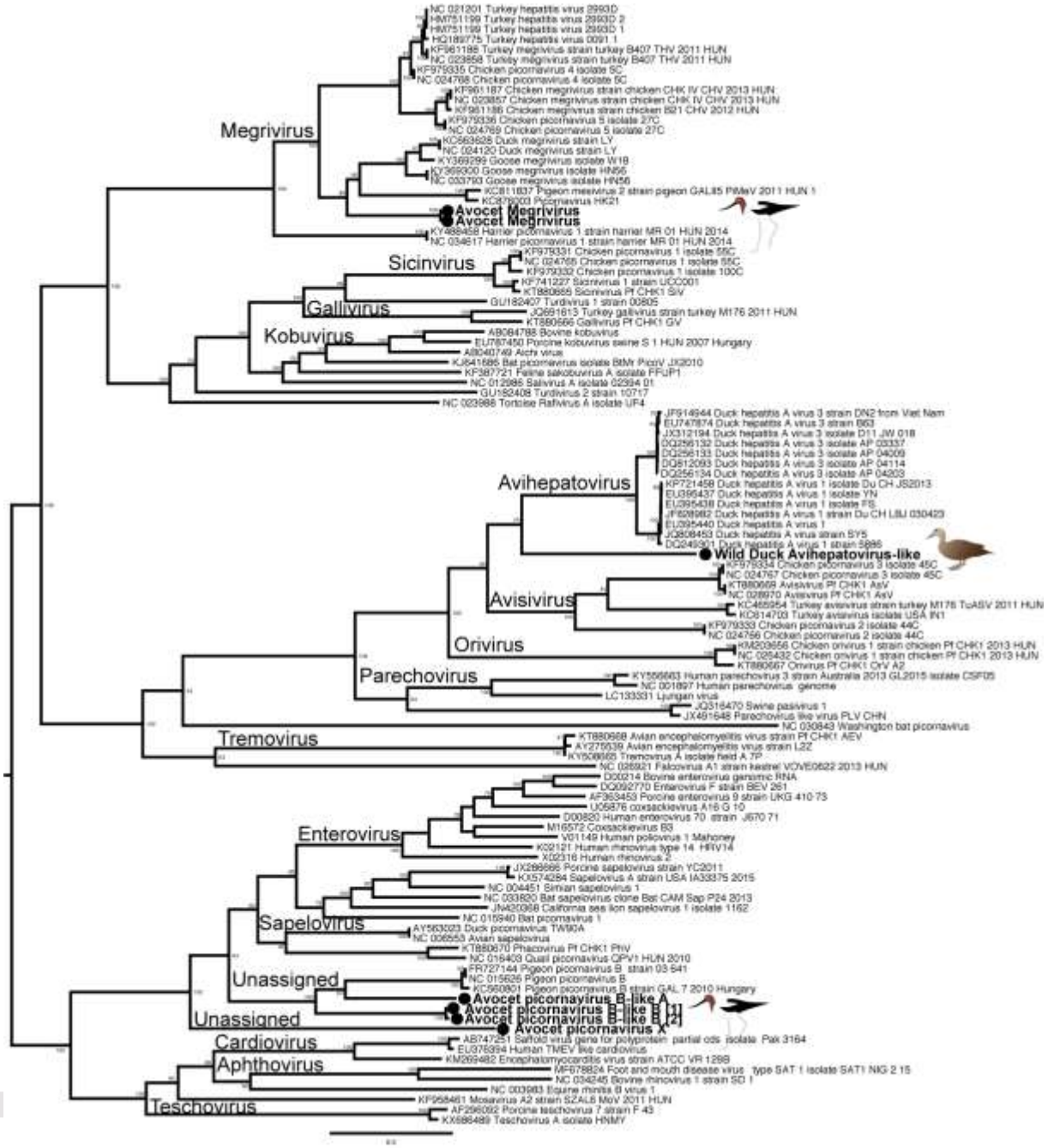

Fig 6. Phylogeny of the polyprotein, containing the RdRp, of the *Picornaviridae*. Reference viruses are those from (Boros *et al.* 2016) and the tree was midpoint rooted for clarity only. The scale bar indicates the number of amino acid substitutions per site. Bootstrap values >70% are shown for key nodes. Viruses described in this study are marked with a filled circle. The tree is mid-point rooted for clarity only.

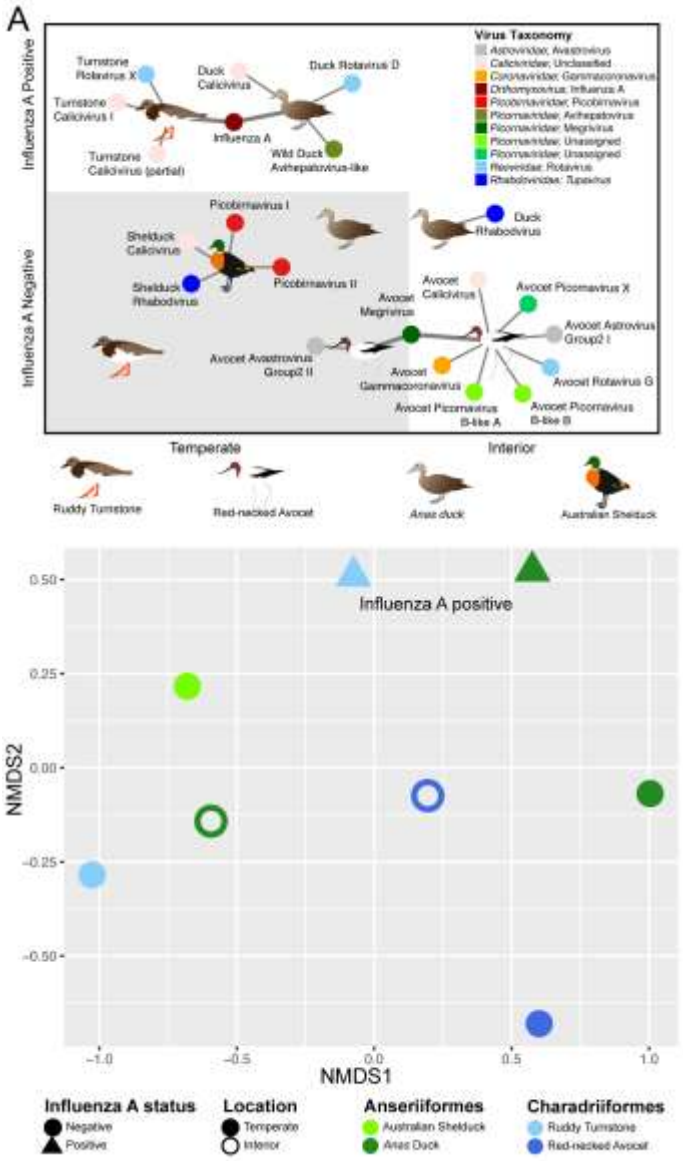

Fig 7. Composition of avian viromes. (a) Bipartite network illustrating species for which complete viral genomes found in each library. Each library is represented as a central node, with a pictogram of the avian species, surrounded by each viral species. Where no complete viral genomes were revealed, the pictogram is shown with no viruses. Where two libraries share a virus species the networks between the two libraries are linked, and the edges are thicker for aesthetic purposes. Placement of libraries are arranged by influenza A status on the y-axis, and location on the x-axis. Virus colour corresponds to virus taxonomy. A list of viruses from each library is presented in Table S1, and phylogenetic trees for each virus family and species can be found in (Fig 2-6, Figs S3-S8). (b) Non-metric multidimensional scaling (NMDS) plot (applying the Bray Curtis dissimilarity matrix) for viral abundance and virus family diversity. Colour, shape and fill correspond to host species, influenza A infection status and location, respectively. For increased clarity, influenza A positive libraries are indicated. Additional NMDS plots where data are analysed with and without IAV reads at both the viral genus and family level are found in Figs S12-S13.

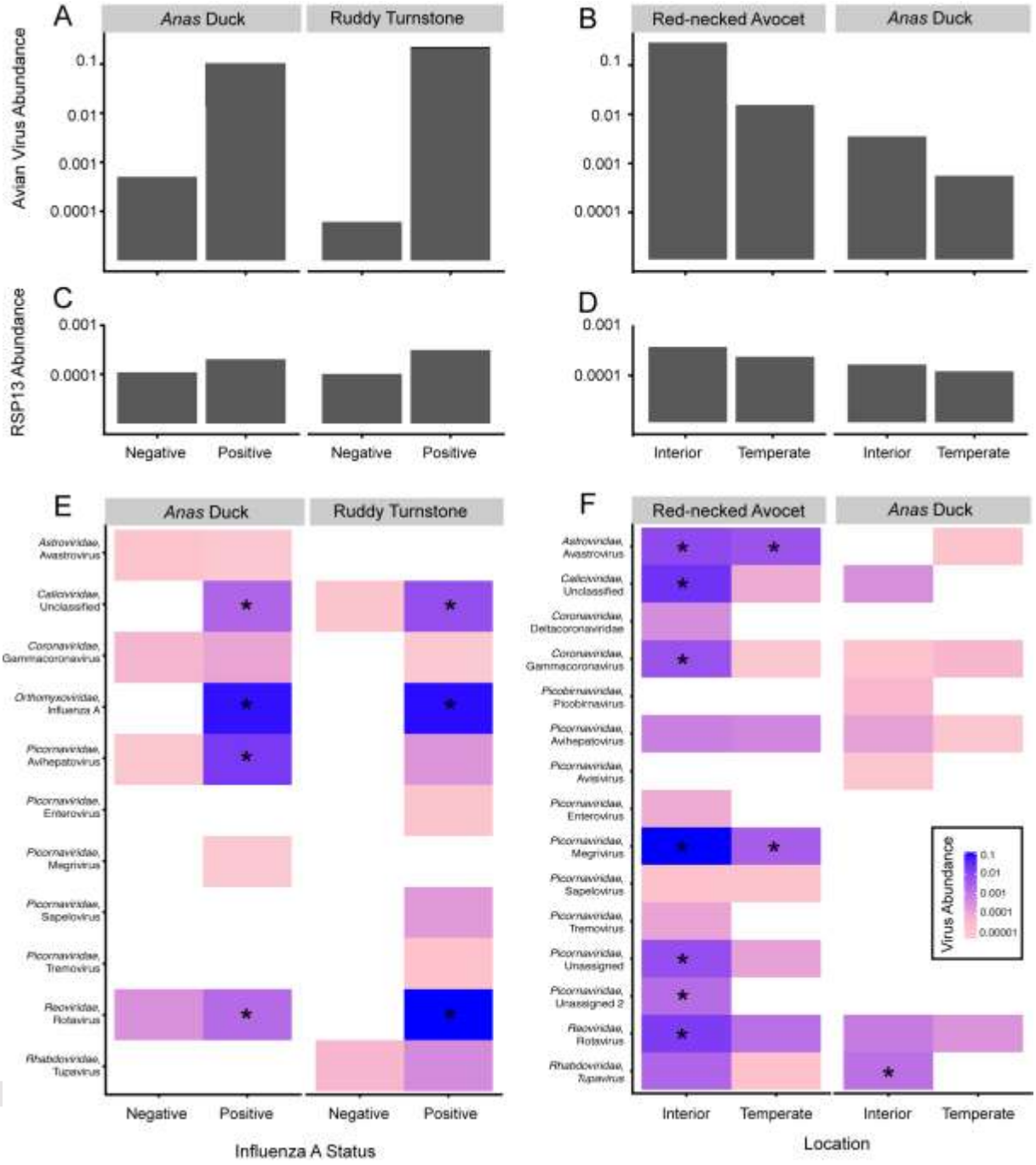

Fig 8. Influenza A status and location are associated with differences in viral abundance and diversity. A,C,E correspond to the influenza A virus infection status, while B, D, F correspond to location. (A, B) Avian viral abundance in libraries in grey, and in (A) abundance of IAV is indicated in black. (C, D) Abundance of host reference gene RSP13. (E, F) Heatmap illustrating viral diversity, at the genus level in each library, with colour corresponding to viral abundance. Blue and purple correspond to viruses with high abundance and pink correspond to viruses with low abundance. Asterisks indicate cases in which at least one complete viral genome was obtained.