



Fecal Viral Diversity of Captive and Wild Tasmanian Devils Characterized Using Virion-Enriched Metagenomics and Metatranscriptomics

Rowena Chong,^a Mang Shi,^{b,c} Catherine E. Grueber,^{a,d}  Edward C. Holmes,^{b,c} Carolyn J. Hogg,^a Katherine Belov,^a Vanessa R. Barrs^{b,e}

^aSchool of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia

^bMarie Bashir Institute for Infectious Diseases and Biosecurity, Sydney Medical School, University of Sydney, Sydney, New South Wales, Australia

^cSchool of Life and Environmental Sciences and Sydney Medical School, Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia

^dSan Diego Zoo Global, San Diego, California, USA

^eSydney School of Veterinary Science, University of Sydney, Sydney, New South Wales, Australia

ABSTRACT The Tasmanian devil is an endangered carnivorous marsupial threatened by devil facial tumor disease (DFTD). While research on DFTD has been extensive, little is known about viruses in devils and whether any are of potential conservation relevance for this endangered species. Using both metagenomics based on virion enrichment and sequence-independent amplification (virion-enriched metagenomics) and metatranscriptomics based on bulk RNA sequencing, we characterized and compared the fecal viromes of captive and wild devils. A total of 54 fecal samples collected from two captive and four wild populations were processed for virome characterization using both approaches. In total, 24 novel marsupial-related viruses, comprising a sapelovirus, astroviruses, rotaviruses, picobirnaviruses, parvoviruses, papillomaviruses, polyomaviruses, and a gammaherpesvirus, were identified, as well as known mammalian pathogens such as rabbit hemorrhagic disease virus 2. Captive devils showed significantly lower viral diversity than wild devils. Comparison of the two virus discovery approaches revealed substantial differences in the number and types of viruses detected, with metatranscriptomics better suited for RNA viruses and virion-enriched metagenomics largely identifying more DNA viruses. Thus, the viral communities revealed by virion-enriched metagenomics and metatranscriptomics were not interchangeable and neither approach was able to detect all viruses present. An integrated approach using both virion-enriched metagenomics and metatranscriptomics constitutes a powerful tool for obtaining a complete overview of both the taxonomic and functional profiles of viral communities within a sample.

IMPORTANCE The Tasmanian devil is an iconic Australian marsupial that has suffered an 80% population decline due to a contagious cancer, devil facial tumor disease, along with other threats. Until now, viral discovery in this species has been confined to one gammaherpesvirus (dasyurid herpesvirus 2 [DaHV-2]), for which captivity was identified as a significant risk factor. Our discovery of 24 novel marsupial-associated RNA and DNA viruses, and that viral diversity is lower in captive than in wild devils, has greatly expanded our knowledge of gut-associated viruses in devils and provides important baseline information that will contribute to the conservation and captive management of this endangered species. Our results also revealed that a combination of virion-enriched metagenomics and metatranscriptomics may be a more comprehensive approach for virome characterization than either method alone. Our results thus provide a springboard for continuous improvements in the way we study complex viral communities.

KEYWORDS carnivore, endangered species, evolutionary biology, marsupial, microbial ecology, microbiome, virus

Citation Chong R, Shi M, Grueber CE, Holmes EC, Hogg CJ, Belov K, Barrs VR. 2019. Fecal viral diversity of captive and wild Tasmanian devils characterized using virion-enriched metagenomics and metatranscriptomics. *J Virol* 93:e00205-19. <https://doi.org/10.1128/JVI.00205-19>.

Editor Susana López, Instituto de Biotecnología/UNAM

Copyright © 2019 Chong et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Vanessa R. Barrs, vanessa.barrs@sydney.edu.au.

Received 16 February 2019

Accepted 21 February 2019

Accepted manuscript posted online 13 March 2019

Published 15 May 2019

The Tasmanian devil (*Sarcophilus harrisii*) is the world's largest extant carnivorous marsupial and is found in the wild only on the island state of Tasmania, Australia. As this animal is predominately a scavenger, its diet largely comprises carrion of mammals, including wallabies, possums, and kangaroos, although it may also consume and digest fish, insects, fruit, and vegetation (1, 2). Listed as endangered, the Tasmanian devil is facing the threat of extinction due to a contagious cancer, devil facial tumor disease (DFTD1 and -2), that has caused drastic (77%) declines in wild devil populations since its discovery in 1996 (3). In an attempt to save the species from extinction, an insurance population was established in 2006 to supplement wild populations at risk of population crashes (4, 5). While extensive research has focused on DFTD, susceptibility to DFTD, and devil genetic diversity (6–10), understanding of other disease threats to devils remains limited. The gut microbiome has been characterized (11), but virological studies are limited to the identification of a single gammaherpesvirus (dasyurid herpesvirus 2 [DaHV-2]), for which captivity was identified as a significant risk factor (12). Characterization of viral diversity among Tasmanian devils is an essential step to improve understanding of host-microbe relationships, and for conservation management of the species. Comparative analysis of marsupial viruses with those from diverse vertebrate hosts, including eutherian mammals, birds, and other vertebrates, will also provide a deeper understanding of the phylogenetic history of the viruses infecting this evolutionary unique group of mammals (13, 14).

The most widely used method for studying viral metagenomics relies on the enrichment of virions and sequence-independent amplification prior to sequencing (15–17). Removal of nonviral genomic host and bacterial nucleic acids is often necessary for the detection of low-titer viruses (17, 18). More recently, the use of RNA sequencing of total non-rRNA from environmental samples gave rise to viral metatranscriptomics, which has been successfully applied to characterize the viromes of diverse invertebrate and vertebrate species (13, 19, 20). To our knowledge, no studies to date have directly compared these two approaches to virome characterization, although doing so would allow us to understand the detection capabilities and biases associated with different nucleic acid extraction and sample treatment methods.

We characterized the fecal virome of wild and captive Tasmanian devil using both metagenomics based on virion enrichment and sequence-independent amplification (here “virion-enriched metagenomics”) and metatranscriptomics based on RNA sequencing of total non-rRNA (here “metatranscriptomics”). Our objectives were (i) to comprehensively characterize the fecal virome of Tasmanian devils, comparing those of wild and captive devils, and (ii) to compare the two virome characterization approaches, highlighting their advantages and potential challenges.

(This article was submitted to an online preprint archive [21].)

RESULTS

Overview of the devil virome. We characterized the fecal virome of six pools consisting of a total of 54 unique fecal samples collected from 54 individual Tasmanian devils from four wild sites and two captive sites (see Materials and Methods and Fig. 1) using both metatranscriptomics and virion-enriched metagenomics approaches. Metatranscriptomic sequencing resulted in 128 to 140 million reads per pool (793,038,436 reads in total), which were assembled *de novo* into 196,919 to 358,327 contigs (Table 1). BLAST analyses of sequence reads from the metatranscriptomic protocol revealed large proportions of reads from *Bacteria* (55.11 to 67.77%) and only 4.32 to 7.88% from *Eukarya*. Mapping reads to the Tasmanian devil genome revealed that 10.75 to 18.99% of reads originated from the host. The percentage of reads related to *Archaea* was less than 0.02%, and the value for viruses was between 0.68 and 1.16% (Fig. 2a).

Virion-enriched metagenomics resulted in 26 to 49 million reads per pool (237,174,236 reads in total), which were assembled *de novo* into 49,813 to 313,354 contigs (Table 1). Compared to metatranscriptomics, the virion-enriched metagenomics protocol resulted in smaller proportions of nonviral nucleic acids (host and bacterial) and hence enriched viral nucleic acids (Fig. 2a). The proportion of viral reads in virion-enriched metagenomics varied



FIG 1 Map of Tasmania, Australia, showing the four wild sampling sites. The number of fecal samples in each location pool is also indicated. The two captive sampling sites are located on mainland Australia and are designated zoo A and zoo B, with 9 and 10 unique fecal samples in each pool, respectively. (Courtesy of A. V. Lee; reproduced with permission.)

from 14.69 to 60.02%, while the proportions of reads mapped to nonviral components were 17 to 49.46% for *Bacteria*, 0.17 to 2.42% for *Eukarya*, 1.29 to 18.67% for the host, and less than 0.01% for *Archaea* (Fig. 2a). For both approaches, a substantial proportion of reads had no significant similarity to other sequences in the databases in GenBank (10.22 to 29.20%).

TABLE 1 Library information of metatranscriptomics and virion-enriched metagenomics in the present study

Location	C ^a /W ^b	Metatranscriptomics		Virion-enriched metagenomics	
		No. of reads	No. of contigs	No. of reads	No. of contigs
Zoo A	C	139,827,866	219,496	26,704,842	281,857
Zoo B	C	128,149,408	196,919	33,002,790	267,536
Maria Island	W	129,783,390	261,483	42,856,850	153,804
Buckbys Road	W	128,589,556	358,327	42,602,752	84,646
wukalina/Mt William National Park	W	129,981,856	264,205	48,911,728	313,154
Stony Head	W	136,706,360	325,769	43,095,274	49,812
Total		793,038,436	1,626,199	1,150,809	

^aC, captive.

^bW, Wild.

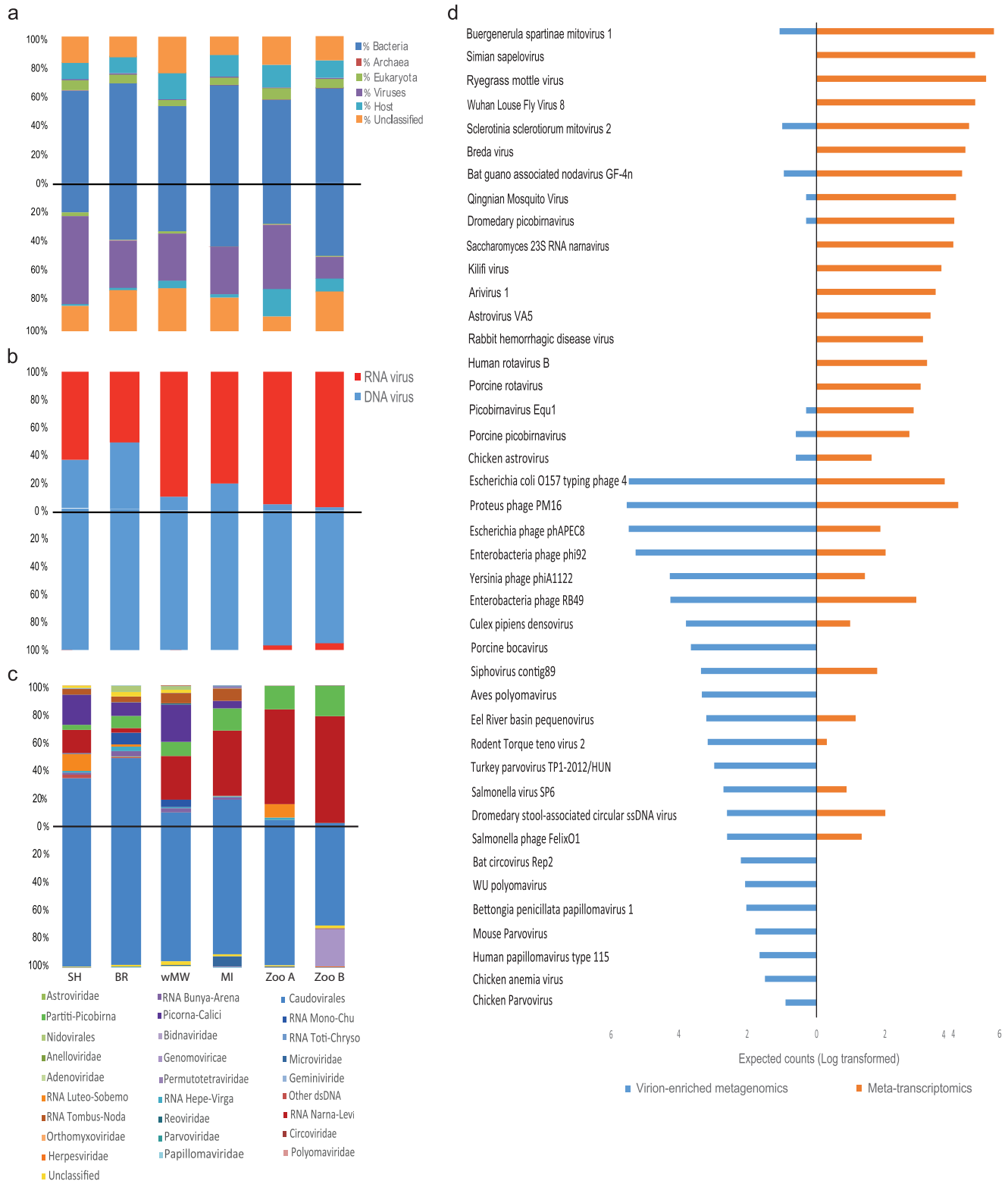


FIG 2 Overview of the Tasmanian devil fecal virome characterized by metatranscriptomics and virion-enriched metagenomics. Sequencing library/sampling site are represented at the bottom of the bar charts by SH for Stony Head, BR for Buckbys Road, wMW for wukalina/Mt William National Park, MI for Maria Island, and Zoo A and Zoo B for the two captive populations. (a) Proportions of all sequence reads in metatranscriptomics (top) and virion-enriched metagenomics (bottom) categorized as belonging to bacteria, eukaryotes, archaea, viruses, and hosts or unidentified. (b) Proportions of RNA and DNA viruses detected by metatranscriptomics (top) and virion-enriched metagenomics (bottom). (c) Virome composition and the proportions of viral groups in metatranscriptomics (top) and virion-enriched metagenomics (bottom). (d) Estimated counts as calculated by RSEM (all six sequencing libraries from both virion-enriched metagenomics and metatranscriptomics combined and log transformed) to a selection of viruses, showing the differences in viruses detected by virion-enriched metagenomics and metatranscriptomics.

Overall, virion-enriched metagenomics and metatranscriptomics differed in the viruses detected as well as the expected counts (transcript abundance) as measured by RSEM analysis (Fig. 2c and d). Viruses from a wide range of viral groups were detected with metatranscriptomics, of which 49.87 to 97.51% had the closest hits to RNA viruses and 2.49 to 50.13% to DNA viruses. Conversely, for virion-enriched metagenomics, >95.54% of the virus-related sequences had the closest hits to DNA viruses, and <5% were identified as RNA viruses (Fig. 2b). Metatranscriptomics revealed high levels of viral diversity across all libraries, the most abundant viral groups detected being *Caudovirales*, *Luteo-Sobemo*, *Narna-Levi*, *Partiti-Picobirna*, *Picorna-Calici*, and *Tombus-Noda* (Fig. 2c). Conversely, virion-enriched metagenomics revealed relatively lower viral diversity across the same libraries; *Caudovirales* dominated the viral reads (69.89 to 99.49%), while viral groups identified at much lower abundances included *Microviridae*, *Circoviridae*, *Genomoviridae*, *Parvoviridae*, *Herpesviridae*, *Polyomaviridae*, and *Papillomaviridae* (Fig. 2c).

Vertebrate viruses detected by metatranscriptomics comprised 0 to 9.41% of the total viral reads. A large proportion of the viral reads belonged to either nonvertebrate eukaryotic viruses (45.08 to 97.51%), including plant viruses, insect viruses, and mycoviruses, or bacteriophage (2.48 to 48.91%) from the families *Siphoviridae*, *Podoviridae*, *Myoviridae*, and *Microviridae*. In the virion-enriched metagenomics data set, the percentage of vertebrate virus reads was also small (0.04 to 0.84%), while bacteriophage and other eukaryotic virus reads ranged between 79.17 and 99.91% and between 0.04 and 19.99%, respectively. Detailed information on all vertebrate viruses identified is presented in Table S1 in the supplemental material.

Detection of viruses previously identified in other mammalian hosts (rabbit hemorrhagic disease virus and torovirus). Rabbit hemorrhagic disease virus (RHDV) is a calicivirus in the genus *Lagovirus* (22). RHDV is used as a biocontrol for rabbits in Australia and causes fatal hepatitis in European rabbits (*Oryctolagus cuniculus*) and some hare species (22). Using metatranscriptomics, we detected genomes with high nucleotide and amino acid similarity (>98%) to RHDV in one of the wild devil metatranscriptomic libraries (Buckbys Road [BR]), with genome coverage of 98.1%. Phylogenetic analysis based on the nucleotide sequences of the major capsid and nonstructural protein genes revealed that the RHDV detected in this study clustered with RHDV variant Gl.2 (also called RHDV2) (Fig. 3a and b), which was first detected in Australia in May 2015 and has since become the dominant circulating variant nationwide. RHDV-specific reverse transcription-PCR (RT-PCR) and sequencing confirmed the presence of RHDV2 in four of the nine devil fecal samples from the BR metatranscriptomics pool. In addition, no rabbit-associated genes were detected during the initial sequence analysis. Additional PCR targeting a short fragment of rabbit mitochondrial DNA (mtDNA) (<300 bp) also did not detect any rabbit DNA in the original fecal samples from BR. Further RT-PCR and sequencing performed on the fecal RNA extractions from the remaining pools confirmed the presence of RHDV in 1 of 10 devils from wukalina/Mt William National Park (wMW), two of nine devils from zoo A, and one of nine from zoo B. One of the four additional RHDV-positive samples, from zoo A, contained rabbit mtDNA as confirmed by PCR.

We also identified the complete viral genome (28,463 bp) of a novel torovirus most closely related (96% nucleotide similarity) to bovine torovirus (Breda virus) in one of the metatranscriptomic libraries (wMW). We determined the full genome structure of the novel torovirus variant, including open reading frame (ORF) 1a and ORF 1b (encoding the two polyproteins pp1a and pp1ab), ORF 2 (encoding the spike protein [S]), ORF 3 (encoding the membrane protein [M]), ORF 4 (encoding the hemagglutinin-esterase protein [HE]), and ORF 5 (encoding the nucleocapsid protein [N]) (23). Based on the phylogenetic analysis of the spike protein amino acid sequence (4,762 amino acids [aa]), clustering of the novel torovirus variant with other toroviruses isolated from cattle in the United States, Japan, and Europe tentatively suggested a bovine origin, although this will need to be confirmed with wider sampling (Fig. 3c).

Detection and characterization of novel marsupial-associated viruses. (i) Picornaviruses. The complete genome (8,015 bp) of a novel virus in *Picornaviridae* was

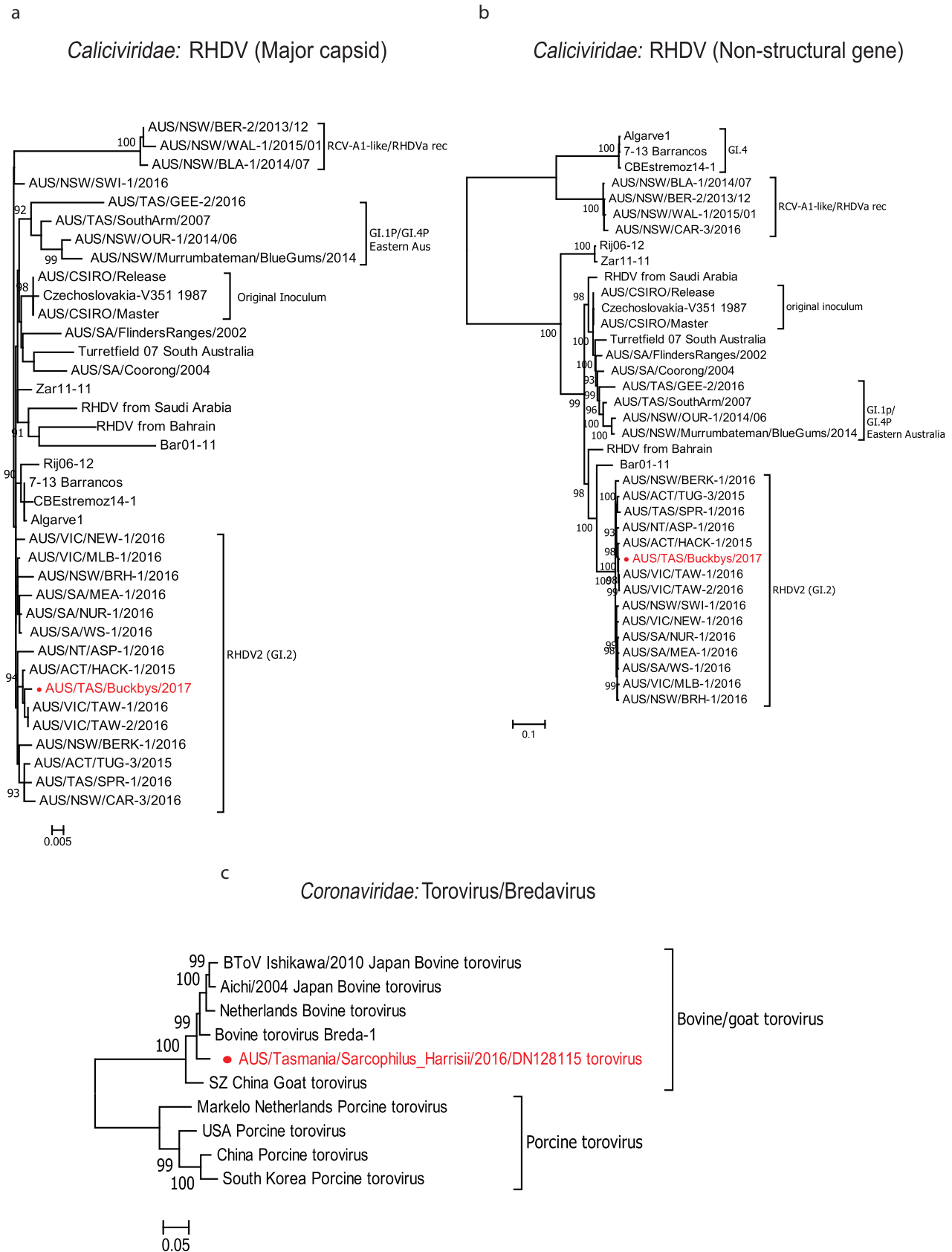


FIG 3 Maximum likelihood phylogenetic trees of viruses detected that were previously identified in other mammalian hosts. (a) Phylogenetic analysis of representative RHDV strains based on the 1,414-nucleotide (nt) sequence of major capsid protein. (b) Phylogenetic analysis of representative RHDV (Continued on next page)

identified in one metatranscriptomic library (wMW). According to the International Committee on Taxonomy of Viruses (ICTV), members of a *Picornavirus* genus should share at least 40% amino acid sequence identity in the polyprotein region (23). The encoded 2,396-aa polyprotein of the virus detected in this study exhibited 45.5% amino acid similarity to simian sapelovirus, placing it in the genus *Sapelovirus*. We have provisionally named this newly identified virus Tasmanian devil-associated sapelovirus. Phylogenetic analysis based on the amino acid sequence of the RNA-dependent RNA polymerase (RdRp) domain showed that Tasmanian devil-associated sapelovirus formed a sister lineage to sapeloviruses identified from eutherian mammals (i.e., porcine and simian sapeloviruses) (Fig. 4a).

(ii) Astroviruses. Astrovirus-related sequences were detected in five of the six metatranscriptomic libraries. We identified one complete and one near-complete genome sequence with 81.4% pairwise nucleotide identity, involving two separate variants of a single astrovirus species, tentatively named Tasmanian devil-associated astrovirus 1 (24). The novel astrovirus identified here has a genome structure typical of other astroviruses, with three putative open reading frames (ORF 1a, ORF 1b, and ORF 2) encoding the protease, RdRp, and capsid, respectively. In addition, we found a ribosomal frameshift motif (AAAAAAC) within the ORF 1a/1b overlap region. Phylogenetic analysis based on the conserved RdRp domain showed that Tasmanian devil-associated astrovirus 1 formed a distinct cluster that is more closely related to astroviruses of mammalian hosts (mamastroviruses) than those of avian hosts (avastroviruses) (Fig. 4b).

(iii) Rotaviruses. Rotavirus sequences were identified in three metatranscriptomic libraries (Stony Head [SH], BR, and wMW). Among them, we identified two segments (3,481 bp and 3,479 bp) encoding rotavirus VP1 (i.e., RdRp). The two RdRp sequences shared >90% nucleotide identity with each other, indicative of two different variants from the same species. In addition, a contig encoding a partial rotavirus VP1 of 282 aa sharing 51% sequence similarity with rotavirus H was also detected in one library (wMW). We named the two viruses Tasmanian devil-associated rotavirus 1 and Tasmanian devil-associated rotavirus 2. The VP1 of Tasmanian devil-associated rotavirus 1 shared the highest amino acid similarity, ~51%, with rotavirus G, while Tasmanian devil-associated rotavirus 2 shared the highest amino acid similarity, 44%, with rotavirus H. Phylogenetic analysis with other rotavirus species based on VP1 suggested that the two Tasmanian devil-associated rotaviruses form a distinct cluster most closely related to rotaviruses from avian and mammalian hosts (Fig. 4c).

(iv) Picobirnaviruses. We detected picobirnavirus sequences that encoded complete and partial viral RdRp (330 to 557 aa) in all four of the six metatranscriptomics libraries from wild devils. However, the picobirnavirus sequences detected in one library (Maria Island [MI]) were too short to be phylogenetically informative and were discarded in the phylogenetic analysis. The novel picobirnaviruses detected in this study are provisionally named Tasmanian devil-associated picobirnaviruses 1, 2, 3, 4, 5, and 6, with two separate variants in Tasmanian devil-associated picobirnaviruses 1 and 5. Phylogenetic analysis based on the RdRp domain of these novel picobirnaviruses showed that they are highly diverse and widely distributed across the phylogeny of this family (Fig. 4d).

(v) Parvoviruses. We identified eight new members of the vertebrate-associated subfamily *Parvovirinae*, including two from the genus *Bocaparvovirus* and six from the recently determined genus *Chapparvovirus*. We recovered partial and near-complete protein sequences sharing ~50% identity to California sea lion bocavirus and porcine bocavirus, respectively. Two bocavirus-related sequences detected in this study shared >97% amino acid sequence similarity, indicating two separate variants of the same

FIG 3 Legend (Continued)

strains based on the 5,890-nt sequence of the nonstructural protein. (c) Phylogenetic analysis based on the 4,762-aa spike protein of the novel torovirus variant detected in this study with other previously identified toroviruses. All trees are mid-point rooted and scaled to either the number of amino acid substitutions or nucleotide substitutions per site based on the nature of alignment. Bootstrap values (>80%) are shown at the key nodes. The viral sequences detected in this study are shown in red in each tree.

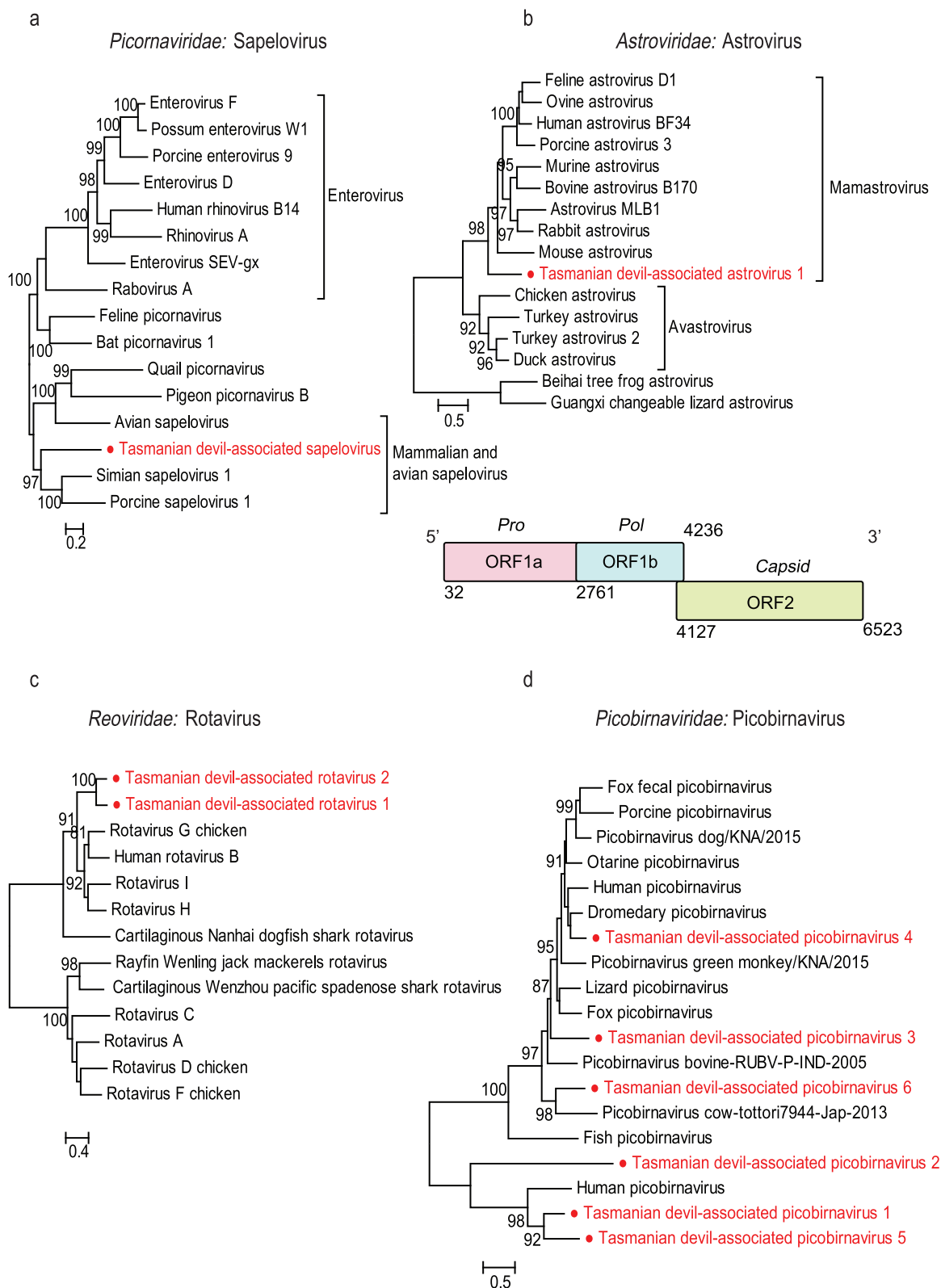


FIG 4 Phylogenetic analyses and genomic structures of the RNA viruses identified in the feces of Tasmanian devils. All phylogenetic analyses were performed based on the amino acid sequence of the RdRp. (a) Tasmanian devil-associated sapelovirus; (b) Tasmanian devil-associated astrovirus; (c) Tasmanian devil-associated rotaviruses 1 and 2; (d) Tasmanian devil-associated picobirnaviruses 1 to 6. For Tasmanian devil-associated astrovirus, for which the whole-genome sequence was obtained, the genomic structure is shown below the corresponding phylogenetic tree. Predicted ORFs of these genomes are labeled with information of the potential protein or protein domain they encode. All trees are mid-point rooted and scaled to the number amino acid substitutions per site. Bootstrap values (>80%) are shown at the key nodes. The newly discovered viruses are shown in red in each phylogenetic tree.

species, provisionally named Tasmanian devil-associated bocavirus 1. A third bocavirus-related sequence was also identified in virion-enriched metagenomics library BR, sharing 71.13% amino acid similarity with Tasmanian devil-associated bocavirus 1 and provisionally named Tasmanian devil-associated bocavirus 2. The six new chapparviruses identified in this study have been provisionally named Tasmanian devil-associated chapparviruses 1 to 6, which shared less than 70% amino acid identity among themselves. Phylogenetic analysis with representative viruses from the *Parvoviridae* family confirmed the clustering of the Tasmanian devil-associated bocaviruses within the diversity of mammalian bocaviruses, although the branching order involving them and other bocaviruses remains unresolved. The chapparviruses identified in this study clustered closely with other chapparviruses, including a recently described parvovirus associated with kidney diseases in mice (25) (Fig. 5a).

(vi) Papillomaviruses. Fragmented genomes of two novel species of papillomavirus were identified in one virion-enriched metagenomics library (MI), among which we retrieved two longer fragments (1,225 and 1,335 bp), both of which encode partial E1 protein, an ATP-dependent DNA helicase required for viral replication (26). The fragments share 64% similarity, suggesting two distinct papillomavirus species, tentatively named Tasmanian devil-associated papillomaviruses 1 and 2. Phylogenetic analysis based on the E1 protein showed that they form a distinct cluster with *Bettongia penicillata* papillomavirus type 1 (BpPV1) isolated from the woylie, a small marsupial species (Fig. 5b and Fig. S3). While the marsupial papillomaviruses are clustered together in the phylogenetic tree, their relationship with viruses identified from eutherian mammals remains unresolved.

(vii) Polyomaviruses. Two novel polyomaviruses were detected in three virion-enriched metagenomics libraries (MI, wMW, and zoo B). We recovered the complete circular genome of 4,894 bp of the tentatively named Tasmanian devil-associated polyomavirus 1 and a partial gene sequence (2,251 bp) of the large T antigen (LTag) protein for the second polyoma-like virus, tentatively named Tasmanian devil-associated polyoma-like virus 2 (Fig. 5c). Phylogenetic analyses revealed strikingly different evolutionary histories for the structural and nonstructural parts of the genome (Fig. 5c and Fig. S4 and S5), indicative of recombination (27). In the LTag phylogeny, Tasmanian devil-associated polyomavirus 1 and Tasmanian devil-associated polyoma-like virus 2 formed a distinct lineage with another marsupial virus, bandicoot papillomatoses carcinomatoses virus type 2 (BPCV-2), which, in turn, clustered with polyomaviruses of avian hosts (Fig. 5c and Fig. S5). In contrast, in the VP1 phylogeny (Fig. S4), the marsupial virus showed no close relationship with the avian viruses. Interestingly, the bandicoot papillomatoses carcinomatoses viruses (BPCV-1 and -2) showed a close relationship to Tasmanian devil-associated polyomavirus 1 only in the LTag region and not the VP1 region.

(viii) Herpesviruses. In one captive virion-enriched metagenomics library (zoo A), we identified 70 contigs matching different regions of a novel herpesvirus genome, provisionally named dasyurid herpesvirus 3 (DaHV-3), which totaled 62,821 bp in length and included partial gene sequences of the DNA polymerase (575 aa), major DNA binding protein (465 aa), helicase (396 aa), glycoproteins M (378 aa) and H (427 aa), and major capsid protein (358 aa), among others. On phylogenetic analysis based on these nonstructural and structural proteins, DaHV-3 clustered with other gammaherpesviruses (*Gammaherpesvirinae*) (Fig. S1 and S2). Further phylogenetic analysis based on the DNA polymerase showed that DaHV-3 forms a distinct lineage most closely related to bovine gammaherpesvirus 6 and ovine gammaherpesvirus 2 (Fig. 5d). The previously characterized dasyurid herpesvirus 2 (DaHV-2) isolated from devils (12) could not be included in the phylogenetic analysis because no sequences were available from the same genomic regions. A BLASTx search of the DNA polymerase showed that DaHV-3 exhibited the greatest amino acid similarity (93.3%) with macropodid herpesvirus 3 (MaHV-3) isolated previously from an Eastern gray Kangaroo (12), whose DNA poly-

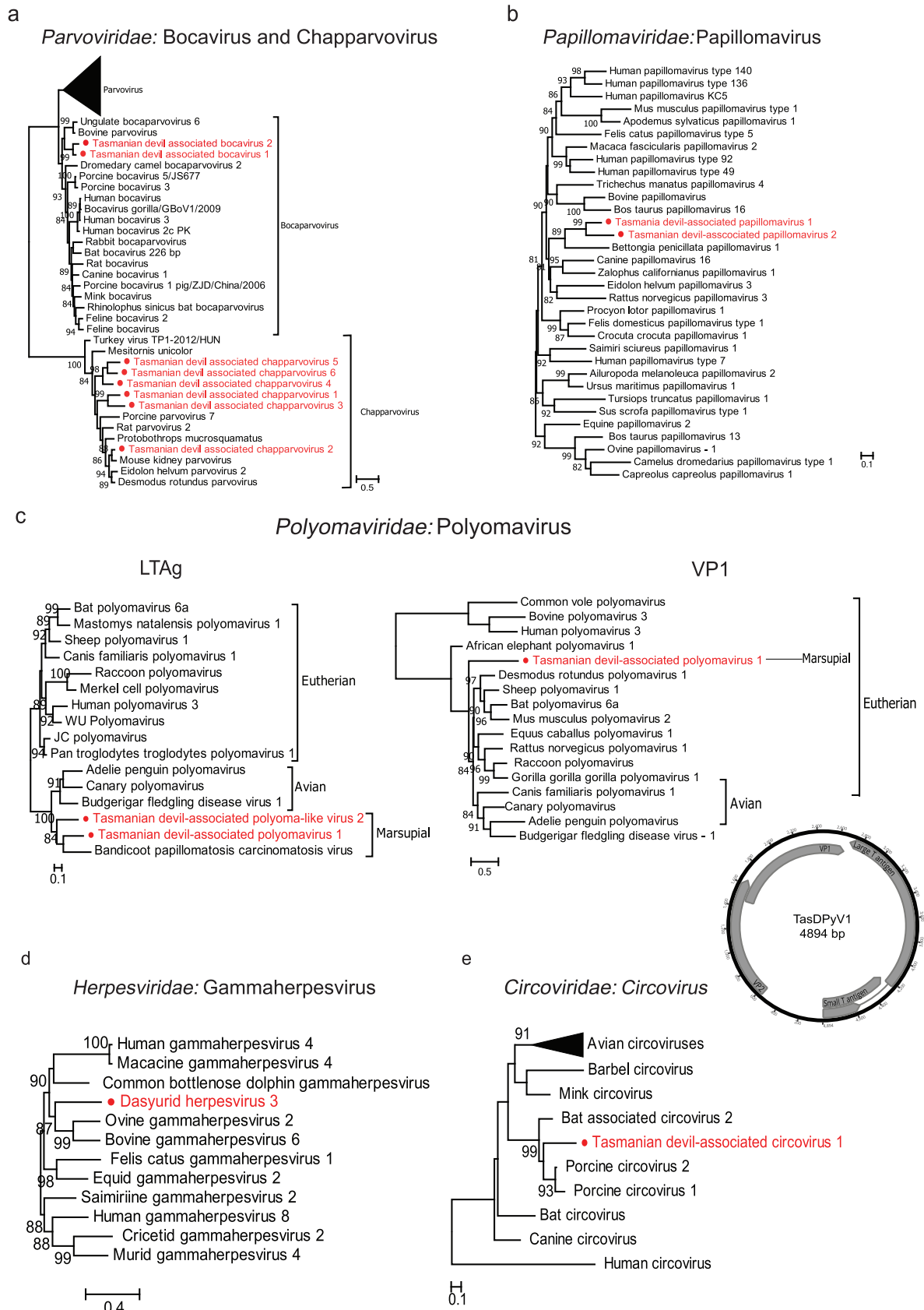


FIG 5 Phylogenetic analyses and genomic structures of the DNA viruses identified in the feces of Tasmanian devils. (a) Tasmanian devil-associated bocaviruses 1 and 2 of *Bocaparvovirus* and Tasmanian devil-associated chapparraviruses 1 to 6 of *Chapparravirus* based on (Continued on next page)

merase amino acid sequence was also too short (<50% of the other representative herpesviruses) to be included in the phylogenetic analysis.

(ix) Circoviruses. We identified circovirus-related sequences and recovered the partial replicase gene sequences (899 bp) in one of the wild devil virion-enriched metagenomes (SH) and tentatively named it Tasmanian devil-associated circovirus. Phylogenetic analysis based on the Rep proteins with representative strains in the *Circovirus* genus suggested that Tasmanian devil-associated circovirus is clustered with circoviruses previously isolated from bats and pigs, sharing the highest sequence identity (62%) with a bat circovirus (GenBank accession number [AIF76281](#)) (Fig. 5e).

Other viruses: plant and insect viruses and bacteriophage. In both the virion-enriched metagenomics and metatranscriptomics analyses, large proportions of viral reads could be attributed to viruses that infect plants, insects, and bacteria, indicating ingestion of arthropods and/or environmental contamination of feces. Bacteriophage sequences from *Caudovirales* were detected in all libraries, comprising >90% of all virion-enriched metagenomic viral reads and up to 48.91% of the metatranscriptomic viral reads. Sequences related to newly identified arthropod viruses, such as Wuhan fly virus and Wuhan mosquito virus, were also detected. Most of the insect viruses detected belong to the *Bunyvirales*, the *Mononegavirales*, and the *Chuviridae*, as well as the DNA virus subfamily *Densovirinae* (*Parvoviridae*). Sequences related to plant and fungal viruses were observed in all libraries, including sobemoviruses, tombusviruses, and mitoviruses.

Comparison between devil populations. Within-library viral diversity as characterized by our metatranscriptomics approach was significantly different between captive and wild populations ($P < 0.05$). In general, captive populations had lower diversity in their fecal viromes than wild populations.

Metatranscriptomics analysis of MI devils displayed a level of overall viral diversity similar to those of both captive populations (zoo A and zoo B) and lower than that found in the other wild populations (Fig. 6a). Conversely, alpha diversity determined from virion-enriched metagenomics data did not differ significantly between libraries (Fig. 6b). Cluster analysis indicated that in metatranscriptomics, the wild and captive devils fell into two distinct clusters, while in virion-enriched metagenomics, BR formed its own cluster and the remaining populations formed a second cluster (Fig. 6).

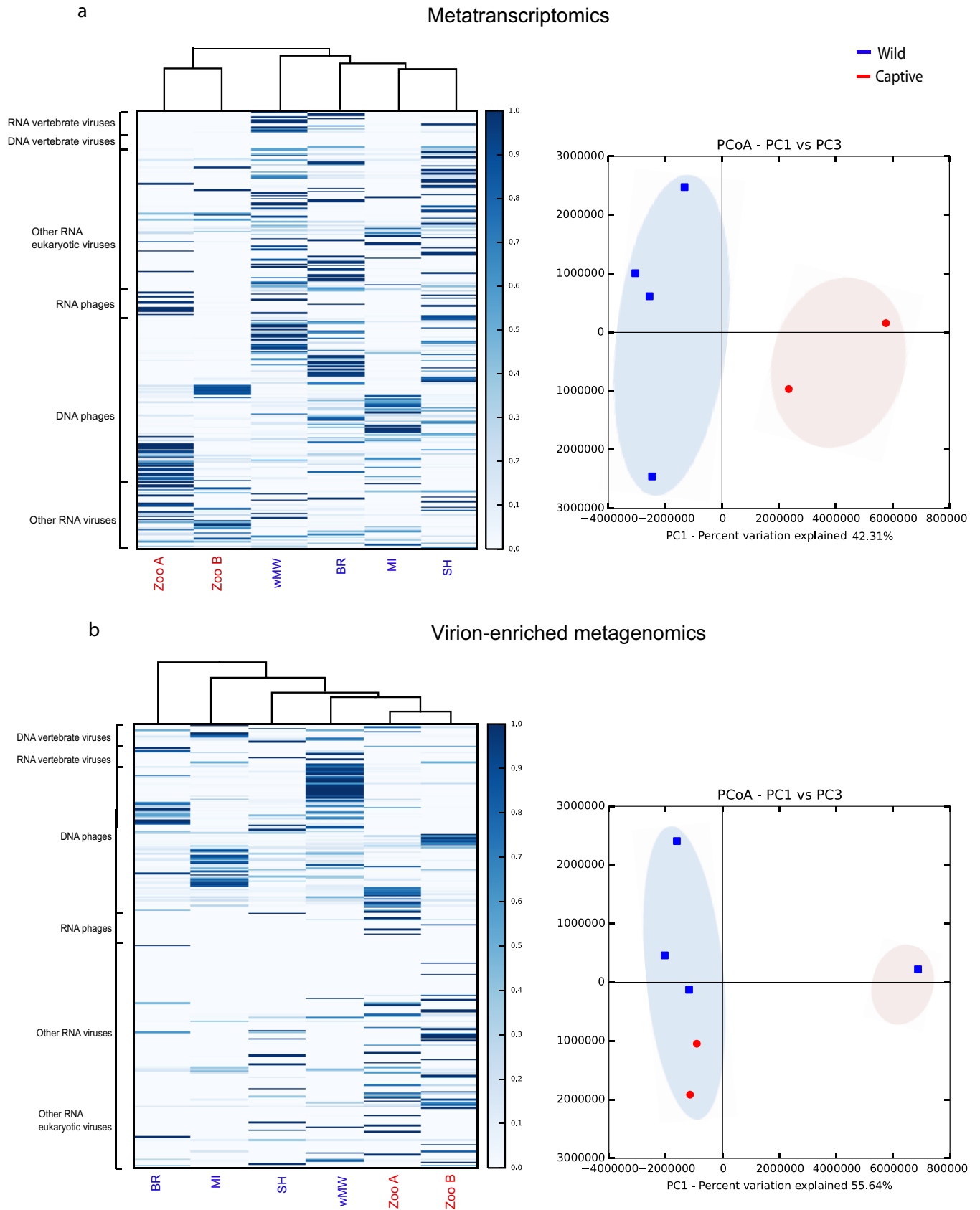
DISCUSSION

Both virion-enriched metagenomics and metatranscriptomics identified a wide diversity of viruses in the feces of Tasmanian devils, including vertebrate viruses, bacteriophages, and other eukaryotic viruses. Overall, we detected sequences related to 26 vertebrate viruses, including 24 novel marsupial-related viruses, including a sapelovirus, an astrovirus, rotaviruses, picobirnaviruses, bocaviruses, chapparvoviruses, papillomaviruses, polyomaviruses, and a gammaherpesvirus, as well as two known mammalian viruses, RHDV2 and torovirus.

However, there were also marked differences between the virion-enriched metagenomics and metatranscriptomics approaches. In general, virion-enriched metagenomics largely detected DNA viruses, while metatranscriptomics detected both DNA and RNA viruses, although the DNA viruses detected were limited to those with relatively high abundance (Fig. 2d). A high abundance level is often indicative of an active infection, during which DNA viruses are transcribed into RNA intermediates detected readily by RNA sequencing (28). Conversely, RNA viruses identified in metatranscrip-

FIG 5 Legend (Continued)

the amino acid sequences of the NS1 protein. (b) Tasmanian devil-associated papillomaviruses 1 and 2 based on the amino acids of the E1 protein. (c) Tasmanian devil-associated polyomavirus 1 and polyoma-like virus 2 based on the amino acid sequences of the LTA_g and VP1 proteins. For Tasmanian devil-associated polyomavirus 1, the whole-genome sequence was obtained and the genomic structure is shown below the VP1 tree. Predicted ORFs of these genomes are labeled with information of the potential protein or protein domain they encode. (d) Dasyurid herpesvirus 3 based on the amino acid sequence of the DNA polymerase. (e) Tasmanian devil-associated circovirus 1 based on the amino acid sequence of the replicase protein. All trees are mid-point rooted and scaled to the number of amino acid substitutions per site. Bootstrap values (>80%) are shown at the key nodes. The newly discovered viruses are shown in red.



tomics were rarely detected in virion-enriched metagenomics, even if they were highly abundant based on the RSEM estimated counts.

The use of virion enrichment and sequence-independent amplification in our metagenomics approach increased the number of viral reads in each library. However, the viral compositions of all six virion-enriched metagenomes were highly skewed toward DNA viruses, particularly bacteriophage from the order *Caudovirales*. Similarly, in a previous study comparing various enrichment methods, bacteriophage accounted for >80% of all reads in all of the enrichment methods tested but <5% when no enrichment steps were incorporated (29). Despite being able to substantially increase the total number of viral reads in the metagenomes, sequence-independent amplification is bias prone, resulting in fewer viruses detected and lower genome coverage due to preferential amplification of certain sequences (30–32). However, the overrepresentation of bacteriophage here may partly be attributed to the fact that they make up the bulk of the gut virobiota, which is dominated by bacteria (33, 34). Regardless of its known bias (29, 30), virion-enriched metagenomics still holds merit for use in virome characterization due to its ability to identify low-abundance DNA viruses, which is especially relevant for dormant or nonactive viruses.

In contrast to virion-enriched metagenomics, metatranscriptomics is nonvirus specific and reveals the entire transcriptome within a sample (20, 35). Since there is no virion enrichment and less sample processing, the likelihood of biased detection is plausibly reduced in metatranscriptomics. In this study, the proportion of viral reads sequenced by metatranscriptomics was <2% per library, but the numbers of viral groups detected were significantly higher than those detected in virion-enriched metagenomics, which included both RNA and DNA viruses. Importantly, then, our results show that the taxonomic compositions of viral communities as revealed by virion-enriched metagenomics and metatranscriptomics were not interchangeable and neither approach was able to detect all viruses present. However, these two approaches were complementary, and an integrated approach using both virion-enriched metagenomics and metatranscriptomics will be a powerful tool for obtaining a complete overview of both the taxonomic and functional profiles of viral communities in a sample.

Ecological analysis of virome composition and diversity revealed significant differences between captive and wild devil populations, especially using metatranscriptomics. Both captive populations displayed lower levels of viral diversity than the wild populations. This loss of diversity is consistent with changes previously observed in the gut bacteriome, where captive devils also exhibited lower bacterial diversity compared to wild devils (11). Changes in lifestyle and diet that occur in captivity likely impact the virome, which is similarly dynamic. Interestingly, Maria Island devils had viromes more similar to those of the captive populations. Maria Island, a 115-km² island off the east coast of Tasmania (Fig. 1), is home to approximately 100 free-ranging devils. Two factors may have contributed to the lower viral diversity observed on Maria Island. First, due to its isolation from mainland Tasmania, animal movements or immigrations are limited to only marine or bird species, and nonendemic species diversity is lower. Thus, there may be limited introduction of viruses to the island. Second, captive-born devils may be more likely to have a “captive-type” virome, that is, lower viral diversity than wild devils. Indeed, some devils included in this study are captive-born animals recently translocated to Maria Island. As samples were pooled prior to sequencing, it was not possible to distinguish between viromes of captive-born and wild-born devils.

While some of the viruses identified in this study come from families that include important pathogens, their pathogenic potential in devils is unclear. It is also important to note that some of these viruses may in fact be dietary viruses, with no active replication in devils, or may occur naturally as part of the normal gut flora of Tasmanian devils. For example, in some areas of Tasmania, rabbits are common and likely ingested by wild devils, which might explain the presence of RHDV2 in wild devils. They are also regularly fed to devils in captivity. Feeding records provided by the two zoos in this study confirmed the feeding of rabbits to some of the sampled devils around the time

of fecal collection. While we did not detect rabbit genes by metatranscriptomics, virion-enriched metagenomics, or PCR in any of the wild devil samples that tested positive for RHDV, we did detect rabbit mtDNA in the feces of one captive devil from zoo A that tested positive for RHDV by PCR. Targeted investigations such as PCR of blood or internal tissues (i.e., liver), *in situ* hybridization, and serological assays are required to determine whether these viruses can actively replicate and cause disease in devils or are simply gut contaminants. Nevertheless, exposure to host-adapted viruses could pose significant health threats, especially for devils that are immunocompromised due to old age, DFTD, or other concurrent diseases (36, 37). Furthermore, even commensal or latent viral infections can be exacerbated or reactivated in immunocompromised hosts (38, 39).

Characterization of the gut virome in healthy devils in this study provides a fundamental baseline for future investigations of diseased animals. Tasmanian devils have low genetic diversity across their genomes and at functionally important loci such as the major histocompatibility complex (MHC) (40–42). This renders them particularly vulnerable to environmental changes, including the emergence of new pathogens, as seen in other species with low genetic diversity (43). For instance, in cheetahs, a coronavirus-associated feline infectious peritonitis outbreak causing mass mortality in a captive breeding colony was linked to the species' extreme genetic monomorphism, particularly at the MHC (44).

Phylogenetic analyses of the newly identified viruses, including divergent members of their respective viral families, has provided insights into the evolutionary history of marsupial-associated viruses relative to viruses of eutherian mammals and other host taxa. Generally, long-term relationships between viruses and hosts are expected for mammalian viruses (13). Strong evidence for this lies in the observation that devil viruses are usually clustered with other marsupial viruses, as a marsupial-specific lineage that is distinct from the eutherian viruses, as observed in herpesvirus, papillomavirus (45, 46), and polyomavirus (Fig. S1 to S5). Furthermore, in several cases the branching order of viruses broadly reflects that of their hosts such that a general codivergence can be inferred. For instance, in the phylogenies of *Picornaviridae* and *Astroviridae*, the Tasmanian devil-associated (marsupial) viruses formed a sister clade to eutherian viruses, which, in turn, are sister to avian viruses, consistent with the evolutionary history of the host. Although such a relationship is not observed in every virus phylogeny, a deep divergence between eutherian and marsupial viruses is typical of our data set. This observation indicates that the timescale of virus evolution is very likely to reflect that of the hosts (13).

The gut virome is increasingly recognized as an integral component of the gut microbiome, and studies of the devil virome will continue to shine light on the biology and health of this iconic endangered species. For example, bacteriophage, which appear to dominate the devil's fecal virome, can contribute to host health by maintaining the diversity and structure of the gut bacteriome through direct interactions with the bacterial communities. While the functions of bacteriophage on devil health remain to be determined, future studies will be able to exploit the extensive microbiomic data that are now available to answer important questions about host-microbe relationship between devils and their microbiome (11).

In sum, our identification of a broad array of vertebrate- and marsupial-specific viruses in devils provides potential candidate viruses for future disease surveillance as part of the broader conservation management of devils once the pathogenic potential of these viruses has been elucidated.

MATERIALS AND METHODS

Sample collection. Fecal samples were collected from healthy wild Tasmanian devils between September 2016 and June 2017 from four locations in Tasmania (Fig. 1)—Stony Head (SH), Buckbys Road (BR), Maria Island (MI), and Wukalina/Mt William National Park (wMW)—and from captive devils at two Australian mainland zoos in June and July 2017 (zoo A and zoo B). Devils were trapped overnight during routine monitoring by Save the Tasmanian Devil Program staff (47). Fresh fecal samples were collected directly from either the animal or the traps or restraint bags during processing of animals. All samples

were stored in liquid nitrogen or a portable -80°C freezer (Stirling Ultracold, Global Cooling Inc.) immediately after collection. After arriving at the laboratory, samples were separated into two aliquots to be used in the extraction of total RNA for metatranscriptomics and enrichment of virions for virion-enriched metagenomics.

Metatranscriptomics: total RNA extraction, library preparation, and sequencing. Samples were disrupted and homogenized in $600\ \mu\text{l}$ of lysis buffer with 1.44-mm ceramic beads using a Bead Ruptor homogenizer (Omni International) at $5\ \text{m}\cdot\text{s}^{-1}$ for 5 min. Total RNA was isolated using the Qiagen RNeasy Plus minikit following the manufacturer's instructions. Extracted RNAs were pooled based on their source locations at equal mass amounts, with each pool containing 6 to 10 samples. Prior to library preparation, RNA pools were depleted of host and bacteria rRNA using a Ribo-Zero-Gold (epidemiology) kit (Illumina). Use of rRNA depletion instead of poly(A) enrichment ensures the retention of RNA with and without poly(A) tails. Sequencing libraries were constructed using a TruSeq total RNA library preparation kit (Illumina), and paired-end (75 bp) sequencing of each library was performed on a NextSeq500 HO platform (Illumina) at the Ramaciotti Centre for Genomics (Sydney, Australia).

Virion-enriched metagenomics. (i) Virion enrichment and nucleic acid extraction. A second aliquot from each fecal sample was processed for the virion-enriched metagenomics approach, as described previously, with minor modifications (17).

Fecal suspensions (10%) were homogenized for 1 min using the Bead Ruptor homogenizer (Omni International) at $5\ \text{m}\cdot\text{s}^{-1}$ and centrifuged at $15,000 \times g$ for 3 min (Hitachi centrifuge, type CT15E; T15A62 fixed-angle rotor). Resulting supernatants were filtered through $0.45\text{-}\mu\text{m}$ membrane filters (Corning), and filtrates were treated with a cocktail of nucleases at 37°C for 2 h. Viral DNA and RNA were then simultaneously extracted using the QIAamp viral RNA minikit (Qiagen) (17).

(ii) Random amplification. Extracted viral nucleic acids were pooled as for the total RNA preparations for metatranscriptomics described above. Pooled extractions were subjected to first- and second-strand synthesis and random PCR amplification for 22 cycles using a complete whole-transcriptome amplification (WTA2) kit (Sigma-Aldrich) (17). WTA2 PCR products were then purified using Agencourt AMPure XP beads (Beckman Coulter) prior to library preparation and sequencing.

(iii) Library preparation and Illumina sequencing. Sequencing libraries were constructed using the Nextera XT DNA library preparation kit (Illumina) according to the manufacturer's instructions, with modifications as described in reference 17. Paired-end (100 bp) sequencing of each library was performed on a HiSeq2500 platform (Illumina) at the Ramaciotti Centre for Genomics.

Assembly and annotation. Sequencing reads were demultiplexed and quality trimmed with Trimmomatic (48) and assembled *de novo* using Trinity (49). Resulting contigs were compared against the nonredundant nucleotide and protein databases on GenBank using BLASTN and BLASTX, respectively, with an E value cutoff at $1\text{E}-5$. BLASTX searching was also conducted against a bespoke database containing all viral RNA-dependent RNA polymerase (RdRp) protein reference sequences downloaded from GenBank. Taxonomic information at the domain level (i.e., *Eukarya*, *Bacteria*, or *Archaea*), as well as for viruses, was assigned based first on the BLASTN results and then on the BLASTX results. Potential virus-related sequences were further categorized into families and orders based on their genetic similarity to their closest relatives and/or their phylogenetic positions. Similarly, assignment of viruses to the broader groups of their hosts (i.e., *Eukarya*, *Bacteria*, or *Archaea*) was based on their phylogenetic relationship to viruses with reliable host information obtained either using experimental or phylogenetic approaches. The genetic identity cutoff for host assignment was 40% based on the most conserved proteins, such as RdRp or DNA polymerases (50). The threshold was set based on the intrafamily diversity of most vertebrate-specific virus families/genera (13). The assignment of vertebrate host was based on phylogenetic analyses, in which a potential devil-associated virus is expected to either cluster within, or form a sister group to, an existing mammalian virus group.

To compare the abundances of transcripts/contigs, we calculated the percentage of total reads in each library. The abundance of host transcripts/contigs was estimated by mapping reads against the Tasmanian devil genome using Bowtie2 (51), whereas those of other organisms, namely, viruses, bacteria, archaea, and nonhost eukaryotes, were estimated using the RSEM approach (52) implemented in Trinity.

For each virus, the genome sequence was further extended by merging related contigs from the same or different pools. Gaps in the genome were filled either by aligning reads to contigs using Bowtie2 (51) or by RT-PCR and Sanger sequencing. Putative ORFs in viral genomes were predicted by the Geneious 8.1 software (53) and annotated based on similarity to previous published virus genomes.

Phylogenetic analysis. Nucleotide sequences of complete or partial genomes and amino acid sequences from the conserved domain (e.g., RdRp) of the newly characterized viral sequences were aligned with those of reference viruses representative of the diversity of the corresponding virus family or species. Alignment was performed using the E-INS-I algorithm implemented in MAFFT (version 7) (54). The quality of the alignments was subsequently assessed, and all ambiguously aligned regions were removed using TrimAl (version 1.2) (55). Phylogenetic trees of aligned amino acid (all data sets with the exception of RHDV) or nucleotide (RHDV) sequences were then inferred using the maximum likelihood method implemented in PhyML (version 3.0) (56), utilizing the best-fit substitution model and the Subtree Pruning and Regrafting (SPR) branch-swapping algorithm.

Analyses of virome ecology. Viral-abundance tables (Table S2 and Table S3) were generated based on complete or near-complete viral contigs and the percentage of reads mapped to them using Bowtie 2 (51) in each library. QIIME (version 1.9) (57) was used to perform ecological and statistical analysis to compare viromes of different populations. Within-library virotype richness (alpha diversity measured using the *Chao1* metric) and dissimilarity between libraries (beta diversity measured using the Euclidean metric) for both virion-enriched metagenomics and metatranscriptomics were calculated based on levels

TABLE 2 List of primer sets used for PCR confirmation of RHDV2

Primer name	Sequence (5'→3')	Amplicon size	Gene
Derived from assembled contigs			
BR_RHDV1_fwd	CTTGCCAGTCTTCTGGAG	168	RdRp
BR_RHDV1_Rev	ACAGGCAAACAGGTCCAAAC		
BR_RHDV3_fwd	ATCCCGCGCCGTTACTGGGTTCCATTAG	405	RdRp
BR_RHDV3_Rev	GGGTCGGGTTTGGTGGGATCTGGAACA		
RHDV strain-specific multiplex (59)			
Gl.1a-Aus_fwd	GCGTGGCATTGTGCCAGCATC	562	RdRp
Gl.1a-Aus_rev	TGTTGGTGATAAGCCATAATCGCG		
Gl.1c_fwd	AGCAAGACTGTTGACTCAATTTTCG	435	VP60
Gl.1c_rev	AGGCCTGCACAGTCGTAACGTT		
Gl.2_fwd	TTCCCTGGAAGCAGTTCGTC	336	VP60
Gl.2_rev	TGTTGTCTGGTTTATGCCATTTGC		
Gl.1a-K5_fwd	TTTATAGATGTATGCCCGCTCAAC	263	RdRp
Gl.1a-K5_rev	CCGTTCGAGTTCCTTGCCGACG		

of viral abundance. Statistical significance of differences in alpha diversity was evaluated by the Monte Carlo method (999 permutations), with a null hypothesis that diversity is equal in all libraries with a significance threshold of $\alpha = 0.05$. Levels of viral abundance were also used to produce heatmaps and dendrograms from hierarchical clustering. Principal-coordinate analysis (PCoA) was performed on the Euclidean distance matrix as calculated in QIIME, and additional cluster analysis was conducted using K-Means clustering in R (58).

PCR confirmation and Sanger sequencing of RHDV. Contigs with high similarity (>97%) to RHDV were detected in one of the metatranscriptomics libraries. To confirm the detection of RHDV, RT-PCR was performed on individual fecal RNA extractions using the Qiagen OneStep Ahead RT-PCR kit (Qiagen) and primers from a previously validated Australian RHDV strain-specific PCR (59) as well as two additional primer sets manually designed based on the current metatranscriptomics assembled contigs (Table 2). PCR products were separated on 1.5% agarose gels (Bio-Rad Laboratories) in 1× Tris-acetate EDTA and visualized using the SYBR Safe DNA gel stain (Life Technologies). Sanger sequencing of positive PCR products was performed at the Australian Genome Research Facility (Sydney, Australia). In addition, DNA was extracted from fecal samples using the Isolate fecal DNA kit (Bioline, London, UK), and the presence of rabbit DNA was tested using primers targeting a 110-bp region of the *Oryctolagus cuniculus* 12S mitochondrial rRNA gene (60). Rabbit DNA extracted from rabbit liver using the Bioline Isolate II genomic DNA kit (Bioline) was used as a positive control. Cleanup, primer trimming, and sequence analysis of Sanger data were performed using Geneious (53).

Data availability. Raw sequence reads generated in this study are available in NCBI SRA database under BioProject number PRJNA495667 (SRA: SRP165630). Viral genome sequences are available in GenBank (DNA viruses under accession numbers MK513523 to MK513543 and RNA viruses under accession numbers MK521912 to MK521927) and are available in the figshare repository (<https://doi.org/10.6084/m9.figshare.7185146.v2>).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JVI.00205-19>.

SUPPLEMENTAL FILE 1, XLSX file, 0.01 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.03 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.03 MB.

SUPPLEMENTAL FILE 4, PDF file, 1.2 MB.

ACKNOWLEDGMENTS

We thank the field teams of the Save the Tasmanian Devil Program and the University of Sydney who collected fecal samples. We thank N. Mooney for his assistance at the BR site, the two zoos (and their keeping staff) which provided access to fecal samples from the captive devils, J. Mahar for providing rabbit liver tissue, and A. V. Lee for the map in Fig. 1. We also acknowledge the Sydney Informatics Hub and the University of Sydney's high-performance computing cluster Artemis for providing the higher-performance resources that have contributed to the research results reported within this paper. Lastly, we thank the two anonymous reviewers for their valuable feedback on improving the manuscript.

This work was funded by an Australian Research Council (ARC; LP140100508) grant

to K.B., C.E.G., C.J.H., and V.R.B. E.C.H. is funded by an ARC Australian Laureate Fellowship (FL170100022).

R.C., V.R.B., M.S., C.J.H., C.E.G., and K.B. designed the experiments. C.J.H. coordinated sample collection from all sites, with R.C., C.E.G., and C.J.H. participating in collection from the zoos, BR, and SH (respectively). R.C. carried out the extractions and PCRs. R.C., M.S., E.C.H., V.R.B., C.J.H., C.E.G., and K.B. analyzed the data. R.C., M.S., E.C.H., and K.B. wrote the manuscript, and all authors critically reviewed and edited the manuscript.

We declare no competing interests.

REFERENCES

- Pemberton D, Gales S, Bauer B, Gales R, Lazenby B, Medlock K. 2008. The diet of the Tasmanian devil, *Sarcophilus harrisii*, as determined from analysis of scat and stomach contents. *Papers Proc R Soc Tasmania* 142:13–22. <https://doi.org/10.26749/rstpp.142.2.13>.
- Owen D, Pemberton D. 2005. Tasmanian devil: a unique and threatened animal, p 63–77. Allen & Unwin, Crows Nest, NSW, Australia.
- Lazenby BT, Tobler MW, Brown WE, Hawkins CE, Hocking GJ, Hume F, Huxtable S, Iles P, Jones ME, Lawrence C, Thalmann S, Wise P, Williams H, Fox S, Pemberton D. 2018. Density trends and demographic signals uncover the long-term impact of transmissible cancer in Tasmanian devils. *J Appl Ecol* 55:1368–1379. <https://doi.org/10.1111/1365-2664.13088>.
- Hogg C, Lee A, Srb C, Hibbard C. 2017. Metapopulation management of an endangered species with limited genetic diversity in the presence of disease: the Tasmanian devil *Sarcophilus harrisii*. *Int Zoo Yearb* 51: 137–153. <https://doi.org/10.1111/izy.12144>.
- Jones ME, Jarman PJ, Lees CM, Hesterman H, Hamede RK, Mooney NJ, Mann D, Pukk CE, Bergfeld J, McCallum H. 2007. Conservation management of Tasmanian devils in the context of an emerging, extinction-threatening disease: devil facial tumor disease. *Ecohealth* 4:326–337. <https://doi.org/10.1007/s10393-007-0120-6>.
- Loh R, Bergfeld J, Hayes D, O'hara A, Pyecroft S, Raidal S, Sharpe R. 2006. The pathology of devil facial tumor disease (DFTD) in Tasmanian devils (*Sarcophilus harrisii*). *Vet Pathol* 43:890–895. <https://doi.org/10.1354/vp.43-6-890>.
- McCallum H. 2008. Tasmanian devil facial tumour disease: lessons for conservation biology. *Trends Ecol Evol* 23:631–637. <https://doi.org/10.1016/j.tree.2008.07.001>.
- Kreiss A, Cheng Y, Kimble F, Wells B, Donovan S, Belov K, Woods GM. 2011. Allorecognition in the Tasmanian devil (*Sarcophilus harrisii*), an endangered marsupial species with limited genetic diversity. *PLoS One* 6:e22402. <https://doi.org/10.1371/journal.pone.0022402>.
- Pearse A-M, Swift K. 2006. Allograft theory: transmission of devil facial-tumour disease. *Nature* 439:549. <https://doi.org/10.1038/439549a>.
- Grueber CE, Peel E, Gooley R, Belov K. 2015. Genomic insights into a contagious cancer in Tasmanian devils. *Trends Genet* 31:528–535. <https://doi.org/10.1016/j.tig.2015.05.001>.
- Cheng Y, Fox S, Pemberton D, Hogg C, Papenfuss AT, Belov K. 2015. The Tasmanian devil microbiome—implications for conservation and management. *Microbiome* 3:1.
- Stalder K, Vaz PK, Gilkerson JR, Baker R, Whiteley P, Ficorilli N, Tatarczuch L, Portas T, Skogvold K, Anderson GA, Devlin JM. 2015. Prevalence and clinical significance of herpesvirus infection in populations of Australian marsupials. *PLoS One* 10:e0133807. <https://doi.org/10.1371/journal.pone.0133807>.
- Shi M, Lin X-D, Chen X, Tian J-H, Chen L-J, Li K, Wang W, Eden J-S, Shen J-J, Liu L. 2018. The evolutionary history of vertebrate RNA viruses. *Nature* 556:197–202. <https://doi.org/10.1038/s41586-018-0012-7>.
- Weisbecker V, Beck R. 2015. Marsupial and monotreme evolution and biogeography. In Kliewe A, Hogan L, Johnston S, Murray P (ed), *Marsupials and monotremes: nature's enigmatic mammals*. Nova Science Publishers, New York, NY.
- Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ. 2008. Viral genome sequencing by random priming methods. *BMC Genomics* 9:5. <https://doi.org/10.1186/1471-2164-9-5>.
- Thurber RV. 2011. Methods in viral metagenomics, p 15–24. In de Bruijn FJ (ed), *Handbook of molecular microbial ecology II: metagenomics in different habitats*. Wiley-Blackwell, Hoboken, NJ.
- Conceicao-Neto N, Zeller M, Lefrere H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijssens J. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 5:16532. <https://doi.org/10.1038/srep16532>.
- Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, Moore NE, Ren X, Huang QS, Carter PE, Peacey M. 2014. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J Virol Methods* 195:194–204. <https://doi.org/10.1016/j.jviromet.2013.08.035>.
- 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 EC, Zhang Y-Z. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539. <https://doi.org/10.1038/nature20167>.
- Shi M, Neville P, Nicholson J, Eden J-S, Imrie A, Holmes EC. 2017. High-resolution metatranscriptomics reveals the ecological dynamics of mosquito-associated RNA viruses in western Australia. *J Virol* 91:e00680-17. <https://doi.org/10.1128/JVI.00680-17>.
- Chong R, Shi M, Grueber CE, Holmes EC, Hogg C, Belov K, Barrs VR. 2018. Characterisation of the faecal virome of captive and wild Tasmanian devils using virus-like particles metagenomics and meta-transcriptomics. *bioRxiv* <https://doi.org/10.1101/443457>.
- Abrantes J, Van Der Loo W, Le Pendu J, Esteves PJ. 2012. Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review. *Vet Res* 43:12. <https://doi.org/10.1186/1297-9716-43-12>.
- King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed). 2012. *Virus taxonomy: classification and nomenclature of viruses: ninth report of the International Committee on Taxonomy of Viruses*. Elsevier, San Diego, CA.
- Bosch AG, Krishna NK, Méndez E, Monroe SS, Pantin-Jackwood M, Schultz-Cherry S. 2012. Family Astroviridae, p 953–959. In King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*. Elsevier, San Diego, CA.
- Roediger B, Lee Q, Tikoo S, Cobbin JCA, Henderson JM, Jormakka M, O'Rourke MB, Padula MP, Pinello N, Henry M, Wynne M, Santagostino SF, Brayton CF, Rasmussen L, Lisowski L, Tay SS, Harris DC, Bertram JF, Dowling JP, Bertolino P, Lai JH, Wu W, Bachovchin WW, Wong JJ-L, Gorrell MD, Shaban B, Holmes EC, Jolly CJ, Monette S, Weninger W. 2018. An atypical parvovirus drives chronic tubulointerstitial nephropathy and kidney fibrosis. *Cell* 175:530–543.e24. <https://doi.org/10.1016/j.cell.2018.08.013>.
- Wilson VG, West M, Woytek K, Rangasamy D. 2002. Papillomavirus E1 proteins: form, function, and features. *Virus Genes* 24:275–290. <https://doi.org/10.1023/A:1015336817836>.
- Buck CB, Van Doorslaer K, Peretti A, Geoghegan EM, Tisza MJ, An P, Katz JP, Pipas JM, McBride AA, Camus AC, McDermott AJ, Dill JA, Delwart E, Ng TFF, Farkas K, Austin C, Kraberger S, Davison W, Pastrana DV, Varsani A. 2016. The ancient evolutionary history of polyomaviruses. *PLoS Pathog* 12:e1005574. <https://doi.org/10.1371/journal.ppat.1005574>.
- Shi M, Zhang Y-Z, Holmes EC. 2018. Meta-transcriptomics and the evolutionary biology of RNA viruses. *Virus Res* 243:83–90. <https://doi.org/10.1016/j.virusres.2017.10.016>.
- Kleiner M, Hooper LV, Duerkop BA. 2015. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. *BMC Genomics* 16:7. <https://doi.org/10.1186/s12864-014-1207-4>.
- Li L, Deng X, Mee ET, Collet-Teixeira S, Anderson R, Schepelmann S, Minor PD, Delwart E. 2015. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J Virol Methods* 213: 139–146. <https://doi.org/10.1016/j.jviromet.2014.12.002>.
- Karlsson OE, Belák S, Granberg F. 2013. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metag-

- enomic detection of viruses. *Biosecur Bioterr* 11:5227–5234. <https://doi.org/10.1089/bsp.2013.0008>.
32. Rosseel T, Ozhelvacı O, Freimanis G, Van Borm S. 2015. Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. *J Virol Methods* 222:72–80. <https://doi.org/10.1016/j.jviromet.2015.05.010>.
 33. Łusiak-Szelachowska M, Weber-Dąbrowska B, Jończyk-Matysiak E, Wojciechowska R, Górski A. 2017. Bacteriophages in the gastrointestinal tract and their implications. *Gut Pathog* 9:44. <https://doi.org/10.1186/s13099-017-0196-7>.
 34. Carding SR, Davis N, Hoyles L. 2017. Review article: the human intestinal virome in health and disease. *Aliment Pharmacol Ther* 46:800–815. <https://doi.org/10.1111/apt.14280>.
 35. Marzano S-Y, Nelson BD, Ajayi-Oyetunde O, Bradley CA, Hughes TJ, Hartman GL, Eastburn DM, Domier LL. 2016. Identification of diverse mycoviruses through metatranscriptomics characterization of the viromes of five major fungal plant pathogens. *J Virol* 90:6846–6863. <https://doi.org/10.1128/JVI.00357-16>.
 36. Cheng Y, Heasman K, Peck S, Peel E, Gooley RM, Papenfuss AT, Hogg CJ, Belov K. 2017. Significant decline in anticancer immune capacity during puberty in the Tasmanian devil. *Sci Rep* 7:44716. <https://doi.org/10.1038/srep44716>.
 37. Kreiss A, Tovar C, Obendorf D, Dun K, Woods G. 2011. A murine xenograft model for a transmissible cancer in Tasmanian devils. *Vet Pathol* 48:475–481. <https://doi.org/10.1177/0300985810380398>.
 38. Wiedinger K, Bitsaktsis C, Chang S. 2014. Reactivation of human polyomaviruses in immunocompromised states. *J Neurovirol* 20:1–8. <https://doi.org/10.1007/s13365-014-0234-x>.
 39. Li L, Deng X, Linsuwanon P, Bangsberg D, Bwana MB, Hunt P, Martin JN, Deeks SG, Delwart E. 2013. AIDS alters the commensal plasma virome. *J Virol* 87:10912. <https://doi.org/10.1128/JVI.01839-13>.
 40. Cheng Y, Sanderson C, Jones M, Belov K. 2012. Low MHC class II diversity in the Tasmanian devil (*Sarcophilus harrisii*). *Immunogenetics* 64:525–533. <https://doi.org/10.1007/s00251-012-0614-4>.
 41. Morris K, Austin JJ, Belov K. 2013. Low major histocompatibility complex diversity in the Tasmanian devil predates European settlement and may explain susceptibility to disease epidemics. *Biol Lett* 9:20120900. <https://doi.org/10.1098/rsbl.2012.0900>.
 42. Miller W, Hayes VM, Ratan A, Petersen DC, Wittekindt NE, Miller J, Walenz B, Knight J, Qi J, Zhao F, Wang Q, Bedoya-Reina OC, Katiyar N, Tomsho LP, Kasson LM, Hardie R-A, Woodbridge P, Tindall EA, Bertelsen MF, Dixon D, Pycroft S, Helgen KM, Lesk AM, Pringle TH, Patterson N, Zhang Y, Kreiss A, Woods GM, Jones ME, Schuster SC. 2011. Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil). *Proc Natl Acad Sci U S A* 108:12348–12353. <https://doi.org/10.1073/pnas.1102838108>.
 43. Willi Y, Van Buskirk J, Hoffmann AA. 2006. Limits to the adaptive potential of small populations. *Annu Rev Ecol Evol Syst* 37:433–458. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110145>.
 44. O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler C, Meltzer D, Colly L, Evermann J, Bush M, Wildt DE. 1985. Genetic basis for species vulnerability in the cheetah. *Science* 227:1428–1434. <https://doi.org/10.1126/science.2983425>.
 45. Van Ranst M, Fuse A, Fiten P, Beuken E, Pfister H, Burk RD, Opendakker G. 1992. Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: comparison of the genome organizations. *Virology* 190:587–596. [https://doi.org/10.1016/0042-6822\(92\)90896-W](https://doi.org/10.1016/0042-6822(92)90896-W).
 46. Tachezy R, Duson G, Rector A, Jenson AB, Sundberg JP, Van Ranst M. 2002. Cloning and genomic characterization of *Felis domesticus* papillomavirus type 1. *Virology* 301:313–321. <https://doi.org/10.1006/viro.2002.1566>.
 47. Hawkins CE, Baars C, Hesterman H, Hocking GJ, Jones ME, Lazenby B, Mann D, Mooney N, Pemberton D, Pycroft S, Restani M, Wiersma J. 2006. Emerging disease and population decline of an island endemic, the Tasmanian devil *Sarcophilus harrisii*. *Biol Conserv* 131:307–324. <https://doi.org/10.1016/j.biocon.2006.04.010>.
 48. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 49. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mucelli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644. <https://doi.org/10.1038/nbt.1883>.
 50. Rost B. 1999. Twilight zone of protein sequence alignments. *Protein Eng* 12:85–94. <https://doi.org/10.1093/protein/12.2.85>.
 51. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357. <https://doi.org/10.1038/nmeth.1923>.
 52. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. <https://doi.org/10.1186/1471-2105-12-323>.
 53. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
 54. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
 55. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
 56. Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704. <https://doi.org/10.1080/10635150390235520>.
 57. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenkov T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
 58. R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
 59. Hall RN, Mahar JE, Read AJ, Mourant R, Piper M, Huang N, Strive T. 2018. A strain-specific multiplex RT-PCR for Australian rabbit haemorrhagic disease viruses uncovers a new recombinant virus variant in rabbits and hares. *Transbound Emerg Dis* 65:e444–e456. <https://doi.org/10.1111/tbed.12779>.
 60. Martín I, García T, Fajardo V, Rojas M, Pegels N, Hernández P, González I. 2010. Polymerase chain reaction detection of rabbit DNA in food and animal feed. *World Rabbit Sci* 17:27–36.