Peer-reviewed version available at Viruses 2019, 11, 211; doi:10.3390/v11030211

Article

Identification of a novel Picorna-like virus, Burpengary virus, that is negatively associated with chlamydial disease in the koala

Erin Harvey[1], Danielle Madden[2], Adam Polkinghorne[2], and Edward C. Holmes*[1]

[1] Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Life and Environmental Sciences and Sydney Medical School, The University of Sydney, Sydney, NSW 2006, Australia; erin.harvey@sydney.edu.au (EH); edward.holmes@sydney.edu.au (ECH);

[2] Animal Research Centre, University of the Sunshine Coast, Sippy Downs, QLD 4556, Australia; danielle.madden@research.usc.edu.au (DM); apolking@usc.edu.au (AP);

* Correspondence: edward.holmes@sydney.edu.au; Tel.: +61 2 9351 5591

Abstract: Koalas (*Phascolarctos cinereus*) are native Australian marsupials whose populations are in decline from a range of threats. Infectious diseases caused by the bacterium *Chlamydia pecorum* and other pathogens are of particular concern. We analysed 26 poly-A selected RNA-sequencing libraries from a data set designed to study the immune response of koalas to ocular chlamydial infection. Using virus discovery techniques, we identified the coding-complete genome sequence of a novel picorna-like virus, denoted Burpengary virus, that was most common in south-east Queensland. Notably, abundance measurements of the virus across all 26 libraries revealed an inverse relationship in prevalence with ocular disease in Koalas, suggesting that co-infection between Burpengary virus and *Chlamydia pecorum* is inhibited.

Keywords: Virus discovery, *Chlamydia pecorum*, Koala, Picornavirus, Phylogeny

1. Introduction

The koala (*Phascolarctos cinereus*) is an iconic Australian marsupial species under threat from a number of anthropogenic factors such as loss of habitat due to deforestation and urbanisation, attack from domestic animals such as cats and dogs, and vehicle collisions [1]. It has been estimated that the koala population of south-east Queensland has declined by as much as 80% since the 1990s. Shrinking population sizes and habitat regions are also contributing to the spread of disease. The infectious threat of most concern to the koala is *Chlamydia pecorum*, a bacterium that infects the eyes and urinary tract causing blindness, infertility and death [2]. It has been estimated that up to 100% of some koala populations are now infected with the bacterium, with most wild populations across Australia affected and those from northern Queensland experiencing higher levels of both infection and clinical disease [3, 4].

In addition to *Chlamydia*, koalas are also known to be susceptible to a number of viral infections such as koala retrovirus (KoRV) and two gammaherpes viruses [5, 6]. KoRV has been linked to leukaemia in koalas and is a potential contributing factor to the high rates of *Chlamydia*, such that KoRV might cause immunomodulation and increase susceptibility to chlamydial disease [7]. To determine if koalas suffering ocular disease due to *Chlamydia* might be infected with other viruses, we employed virus discovery techniques based on bulk RNA sequencing [8-10]. These data were

previously collected to analyse the immune response of koalas to ocular infection, particularly that caused by *C. pecorum* [*11*].

2. Materials and Methods

2.1. Analysis of RNA libraries and identification of a novel virus

A total of 26 previously determined poly-A selected RNA-seq libraries from koalas with evidence of gross pathology consistent with koala chlamydiosis (n = 13) and those without (n = 13) were assembled and analysed [11]. All sequencing data is available at NCBI SRA under Bioproject PRJEB26467. To identify novel viruses within the data set, the paired end reads were trimmed using Trimmomatic (v 0.36) [12] and assembled *de novo* using Trinity (v.2.5.1) [13]. A blast search of the assembled contigs was performed against the NCBI nr database using diamond blast (v 0.9.10) [14]. Hits to virus associated proteins were extracted and a subsequent blast analysis was run against the NCBI nt database using BLASTn [15] to determine if any of these contigs showed significant sequence similarity to previously characterised non-viral sequences. Contigs showing sequence similarity to viral proteins exclusively were then checked for complete open reading frames (ORFs) using the ExPASy translate tool (https://web.expasy.org/translate/). The coding-complete genome sequence of a picorna-like virus was identified in the assembled contigs of four libraries, while partial sequences were identified in nine other libraries. Bowtie2 (v.2.2.5) [16] was then used to retrieve full genomes from all other libraries, where possible using the Trinity contig as a reference. The genome structure of the novel virus was predicted based on the single ORF predicted, and a web blast conserved domain search [17] against the conserved domain database (CDD) was used to predict conserved protein regions.

2.2. Evolutionary relationships and viral abundance

To determine the phylogenetic relationship of the novel virus identified here to previously characterised viruses the amino acid sequence of the assembled polyprotein was aligned to other viruses within the family *Picornaviridae*. Accordingly, all *Picornaviridae* polyprotein amino acid sequences available within the RefSeq database as of January 2019 were compiled and aligned to the novel protein sequence using MAFFT (v.7.300) [18] employing the L-INS-I algorithm. TrimAL (v.1.4.1) [19] was then used to remove ambiguously aligned regions. The IQ-tree (v.1.6.1) tool, ModelFinder [20], was used to determine the best-fit model of amino acid substitution - identified as the Le-Gascuel (LG) model. A maximum likelihood tree assuming this model was then inferred using PhyML (v.20150415) [21], with 1000 bootstrap replications used to determine the support for each node within the tree.

The abundance of the virus in each library was determined using Bowtie2 (v.2.2.5) [16] to align the fastq reads to the assembled genomes of the novel picorna-like virus. The abundance of the housekeeping gene GAPDH and KoRV were also measured using the same method with reference sequences of koala GAPDH and KoRV taken from the RefSeq database. The Bowtie2 read counts were converted to reads per kilobase million (RPKM) to normalise across the three references.

Peer-reviewed version available at Viruses 2019, 11, 211; doi:10.3390/v11030211

2.2. C. pecorum abundance

C. pecorum infectious load was assessed in conjunctival swabs from each animal following DNA extraction as previously described [22]. A recently described *C. pecorum* species-specific LAMP assay was modified for qPCR analysis [23]. The F3 forward primer (5' ATCGGGACCTTCTCATCG 3') and B3 reverse primer (5' GCTGTTGTAAGGAAGACTCC 3') amplify a 209 base pair target region of the gene specific to *C. pecorum.* All reactions were run in duplicate and were carried out on a Rotor Gene Q real-time PCR machine in a 20 µL reaction volume. The reaction mix consisted of 10µL QuantiTect mastermix containing SYBR Green I chemiluminescent dye, 1µL of 10µM each forward and reverse primer, 3 µL dH2O and 5 µL sample DNA template. Cycling conditions included an initial denaturation of 94°C for 10 minutes, with 35 cycles of denaturation (94°C for 15 seconds), annealing (57°C for 30 seconds) and extension (72°C for 25 seconds). Negative (dH2O) and positive (cultured *C. pecorum* DNA) controls were included in each run. Infection loads were expressed as a number of gene target copies/µL.

3. Results

3.1. Data selection

A total of 26 sequencing libraries were chosen from a larger data set originally generated to analyse the immune response of koalas to ocular chlamydial infection [11] (those libraries excluded from our analysis had incomplete information on *C. pecorum* status and geographic location). These RNA-sequencing libraries included samples from 26 individual koalas in varying stages of ocular health grouped into five categories based on disease state and presence or absence of *C. pecorum* following qPCR screening. Specifically, (i) H1 indicates individuals with no signs of ocular disease and the absence of *C. pecorum*; (ii) H2 indicates individuals with no signs of disease but with *C. pecorum* detected by PCR; (iii) G1 are individuals with acute active disease but an absence of symptoms of chronic ocular disease; (iv) G2 indicates individuals with chronic, active disease; and (v) G3 are individuals with chronic, inactive disease with little or no discharge but evidence of corneal scarring.

3.2. Read assembly and analysis

Between 62,555,754 and 132,486,302 paired end reads were generated for each library, which we assembled into between 55,840 and 174,502 contigs using a *de novo* transcript assembly tool. These contigs were submitted to a series of blast searches to exclude host reads and determine if any viruses were present. As expected, KoRV was found in all 26 libraries, although it is challenging to distinguish exogenous from endogenous copies on these data. No other previously identified viruses were found.

3.3. Virus discovery and characterisation

A single novel positive-sense single-stranded RNA virus, which we refer to as Burpengary virus (named after the locality where it was found at highest prevalence), was identified in 15 libraries and a coding complete genome was assembled. A single ORF was predicted, suggesting a polyprotein organisation, with is consistent with the genome structure of other viruses within the family *Picornaviridae*. Within the predicted polyprotein a number of conserved protein sequences were

identified using a blast conserved domain search consistent with the polyprotein structure of other viruses within the family (Figure 1). Analysis of the predicted coding-complete polyprotein amino acid sequence revealed that Burpengary virus is phylogenetically distinct, but most closely related to members of the *Apthovirus* genus that contains Foot-and-mouth virus, Equine rhinitis virus, Bovine picornavirus and Bovine rhinovirus, although with a marked lack of bootstrap support (Figure 1).

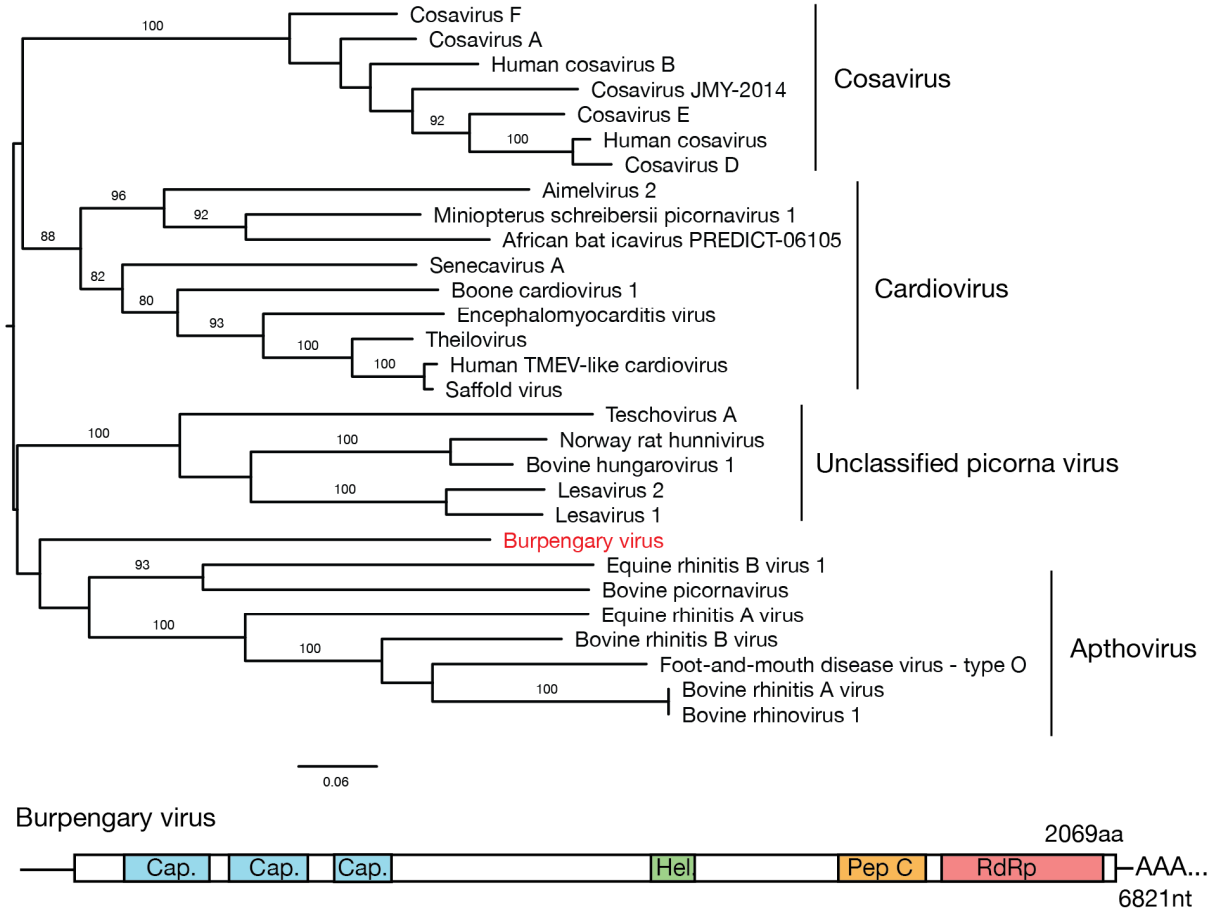

Figure 1. Maximum likelihood phylogenetic tree showing the position of the newly identified Burpengary virus (in red) in relation to other members of the *Picornaviridae*. The tree is midpoint rooted for clarity, and bootstrap support values of over 70% are shown. Genera within the *Picornaviridae* are indicated to the right of the tree. The predicted genome structure of the novel virus is shown below the phylogeny, with the single coding-complete ORF and conserved regions indicated as: Cap - the conserved picornavirus capsid protein domain, Hel - the conserved RNA helicase domain, Pep C - the peptidase C3 cysteine protease domain, and RdRp - the conserved RNA-dependent RNA polymerase.

3.3. Abundance measurement

The abundance of Burpengary virus was measured as reads per kilobase million, and compared to that of koala GAPDH and KoRV to determine if virus abundance was associated with the general abundance of transcripts within each data set. This revealed no association between the abundance of GAPDH or KoRV and Burpengary virus (Figure 2). Importantly, however, when samples were separated by disease status, the presence of Burpengary virus appeared to show an inverse

relationship to signs of ocular disease, being highest in the H1 and H2 samples, and lowest in G1-G3 that designate acute or chronic *C. pecorum* infection (Figure 2). Due to the poly-A selection step being included in the library preparation of these samples, it is impossible to directly compare the abundance of *C. pecorum* with the virus and host genes. The abundance of *C. pecorum* was, however, measured using PCR and gene copies per µL of DNA extract from each conjunctival were recorded for each sample (Figure 2).

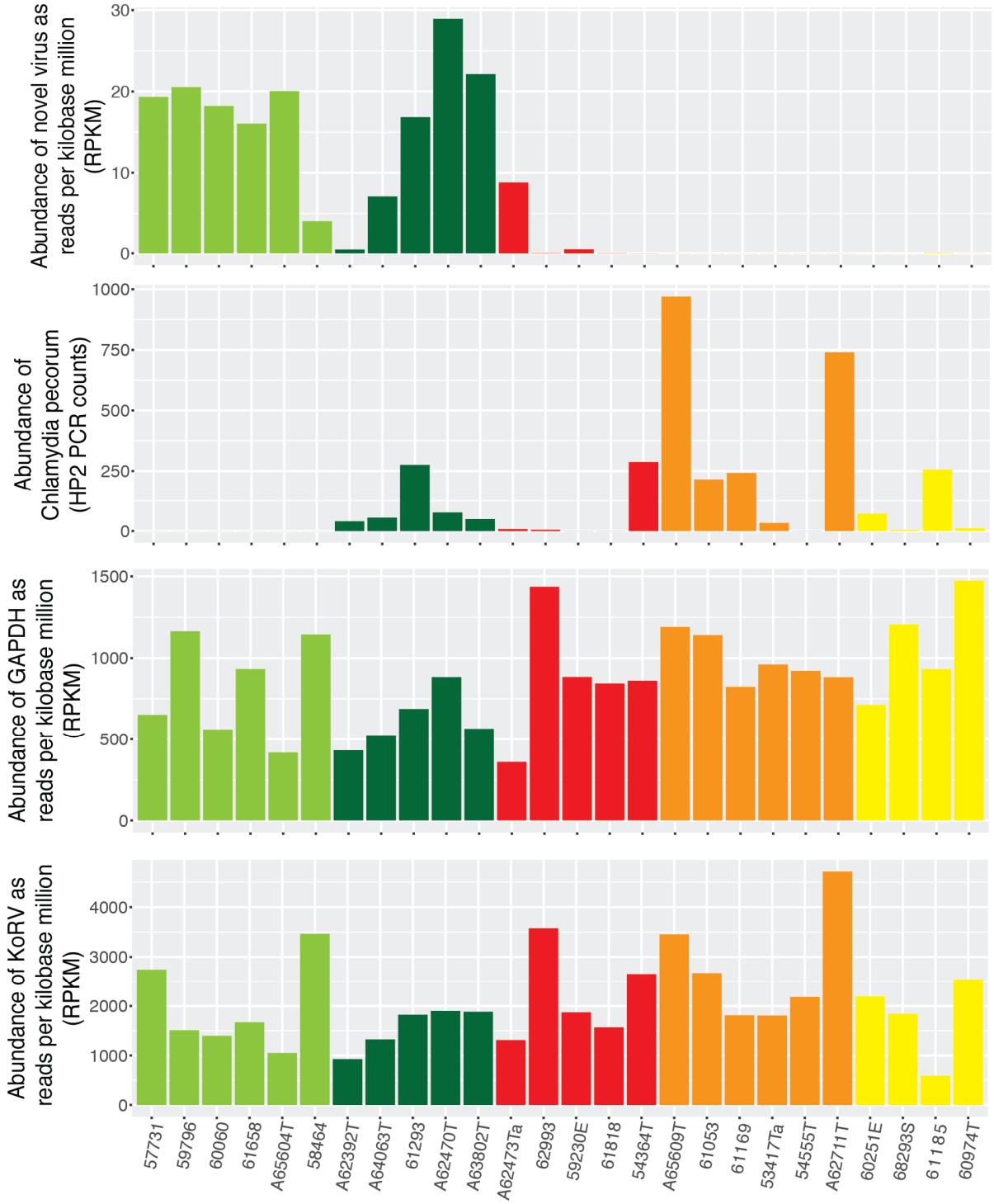

Figure 2. Abundance of Burpengary virus, koala GAPDH, KoRV and *C. pecorum*. Samples are grouped and colour coded by the individual koalas health status as follows: H1 light green, H2 - dark green, G1 - red, G2 - orange and G3 - yellow. The abundance of Burpengary virus, KoRV and GAPDH are measured in reads per kilobase million based on fastq reads, while the abundance of *C. pecorum* was measured by PCR.

3.3. Geographical distrivution of Burpengary virus

Koala samples were collected across south-east Queensland and north-east New South Wales, Australia. To tentatively determine if a particular geographical region was associated with Burpengary virus, the collection locations were mapped and collection points were coloured by virus abundance (Figure 3). This revealed that koalas infected with Burpengary virus were located further north in the collection area, with no virus seen in any samples collected from the border or NSW region. However, this may be an artefact of the distribution of collection locations as most samples were collected around the broader region of south-east Queensland and may be impacted by animal displacement as the koalas tested mainly come from care centres.

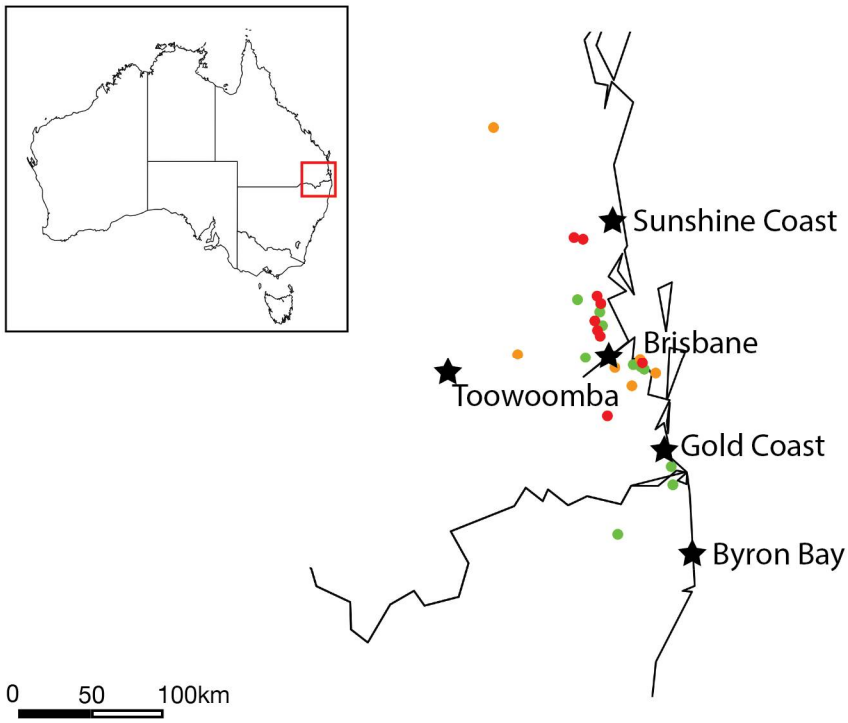

Figure 3. Map showing the koala collection area spanning from north of the Sunshine Coast, Australia, across the New South Wales and Queensland boarder to Byron Bay in the south. Key urban areas are indicated by black stars, while the koala collection locations are colour coded by virus abundance with green points indicating no virus with the sample, yellow indicating low levels of virus (below 0.01% of total reads) and red indicating high abundance of virus (>0.01% of total reads).

4. Discussion

Infectious diseases are a significant driver of morbidity and mortality in wild koalas [1]. The data set analysed here was designed to study the immune response of koalas to ocular chlamydial

infection, and we subsequently analysed 26 poly-A selected sequencing libraries to identify any novel viruses. The technique of eliminating host reads and performing blast searches against a data set of conserved virus protein sequences such as the RNA-dependent RNA polymerase region of RNA viruses has been used previously to identify novel viruses [9, 10, 24, 25]. We implemented this data analysis methodology here and in doing so identified a novel and phylogenetically distinct picornavirus that we have termed Burpengary virus.

As the sequencing libraries were poly-A selected during library preparation, bacterial sequences are necessarily eliminated, such that we were unable to measure the abundance of *C. pecorum* relative to the abundance of Burpengary virus, KoRV or the housekeeping gene GAPDH [11]. This also limits the discovery of novel viruses within the samples as all non-polyadenylated viruses would be excluded from sequencing, and may be a major reason why only a single novel virus was identified. However, Burpengary virus was found in a large proportion of the sequenced samples, with 15 of the 26 libraries containing virus reads.

Interestingly, Burpengary virus was identified almost exclusively in samples collected from healthy koalas showing no signs of ocular chlamydiosis. This tentatively suggests that there is a mechanism, currently unknown, that may prevent the co-infection of ocular tissues with the novel virus and *C. pecorum* when animals develop acute or chronic ocular disease, although this clearly needs verifying with a larger sample size. It is possible that infection with the bacterium may block viral replication in a similar manner to that of *Wolbachia* - a genus of arthropod-associated intracellular bacteria known to block the replication of *Drosophila* and mosquito species with a variety of RNA viruses [*26, 27*], including Dengue virus in *Aedes aegypti* mosquitoes [28, 29]. A perhaps more likely possibility is that chlamydial infection of the mucosal surface by this intracellular bacterial pathogen promotes a strong local immune response, critical to chlamydial infection clearance but also to the promotion of disease [30]. This may affect the presence of other organisms in the conjunctiva. Indeed, although conflicting evidence exists, studies of human trachoma have reported that chlamydial disease is associated with a decrease in microbial richness and diversity compared to healthy controls [31, 32]. Innate immune responses form a key part of the host response to chlamydial infection and, although only speculation, it is possible that upregulation of these processes at the mucosal surface may inadvertently impact on infection by other organisms such as the novel virus detected in this study [30].

These results highlight our limited knowledge of the diversity of novel microorganisms present in unique Australian fauna such as the koala. The clinical impact of Burpengary virus on koala health at the individual and population level is unclear, and like many other studies of this marsupial is limited by the logistical challenges associated with sampling these animals in the field. The apparently inverse association with chlamydial ocular disease appears to add another dimension to the complex and multi-factorial nature of chlamydial disease in this host [30]. Future studies will be required to characterise the significance of the relationship between this novel virus and the major pathogen of the koala, *C. pecorum*.

Author Contributions: conceptualization, A.P, E.H. and E.C.H.; methodology, E.H.; laboratory analysis, D.M.; software, E.H.; formal analysis, E.H.; data curation, E.H.; writing—original draft preparation, E.H.; writing—review and editing, E.C.H and A.P.; visualization, E.H.; supervision, E.C.H.; project administration, A.P. and E.C.H.; funding acquisition, E.C.H.

Funding: ECH is funded by an ARC Australian Laureate Fellowship (FL170100022). AP's work on this project was funded by an ARC Discovery Project (DP150101485).

Acknowledgments: The authors acknowledge the collection of conjunctival samples from koalas in this study by wildlife veterinarians at Australia Zoo Wildlife Hospital, Currumbin Wildlife Sanctuary and Moggill Koala Hospital.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Gonzalez-Astudillo, V.; Allavena, R.; McKinnon, A.; Larkin, R.; Henning, J. Decline causes of koalas in south east Queensland, Australia: a 17-year retrospective study of mortality and morbidity. *Sci. Rep*. **2017**, 7, 42587.
2. Speight, K.N.; Polkinghorne, A.; Penn, R.; Boardman ,W.; Timms, P.; Fraser,T.; Johnson, K.; Faull, R.; Bate, S.; Woolford, L. Prevalence and pathologic features of *Chlamydia pecorum* infections in south australian koalas (*phascolarctos cinereus*). *J. Wildl. Dis*. **2016**, 52, 301-306.
3. Kollipara, A.; Polkinghorne, A.; Wan, C.; Kanyoka, P.; Hanger, J.; Loader, J.; Callaghan, J.; Bell, A.; Ellis, W.; Fitzgibbon, S.; et al. Genetic diversity of *Chlamydia pecorum* strains in wild koala locations across Australia and the implications for a recombinant *C. pecorum* major outer membrane protein based vaccine. *Vet. Microbiol*. **2013**, 167, 513-522.
4. Polkinghorne, A.; Hanger, J.; Timms, P. Recent advances in understanding the biology, epidemiology and control of chlamydial infections in koalas. *Vet. Microbiol*. **2013**, 165, 214-223.
5. Vaz, P.K.; Legione, A.R.; Hartley, C.A.; Devlin, J.M. Detection and differentiation of two koala gammaherpesviruses using high resolution melt (HRM) analysis reveals differences in viral prevalence and clinical associations in a large study of free-ranging koalas. *J. Clin. Micro.* **2019**, doi:10.1128/jcm.01478-18:JCM.01478-01418.
6. Denner, J.; Young, P.R. Koala retroviruses: characterization and impact on the life of koalas. *Retrovirology* **2013**, 10, 108.
7. Tarlinton, R.; Meers, J.; Hanger, J.; Young, P. Real-time reverse transcriptase PCR for the endogenous koala retrovirus reveals an association between plasma viral load and neoplastic disease in koalas. *J. Gen. Virol*. **2005**, 86, 783-787.
8. Shi, M.; Lin, XD.; Tian, J.H.; Chen, L.J.; Chen, X.; Li, C.X.; Qin, X.C.; Li, J.; Cao, J.P.; Eden, J.S.; et al. Redefining the invertebrate RNA virosphere. *Nature* **2016**, 540, 539-543.
9. Pettersson, J.H.; Shi, M.; Bohlin, J.; Eldholm, V.; Brynildsrud, O.B.; Paulsen, K.M.; Andreassen, A.; Holmes, E.C. Characterizing the virome of *Ixodes ricinus* ticks from northern Europe. *Sci. Rep*. **2017**, 7, 10870.
10. Geoghegan, J.L.; Pirotta, V.; Harvey, E; Smith, A; Buchmann, J.P; Ostrowski, M; Eden, J.S; Harcourt, R; Holmes, E.C. Virological sampling of inaccessible wildlife with drones. *Viruses* **2018**, 10, 300.
11. Johnson, R.N; O'Meally, D; Chen, Z; Etherington, G.J; Ho, S.Y.W; Nash, W.J; Grueber, C.E; Cheng, Y; Whittington, C.M; Dennison, S; et al.. Adaptation and conservation insights from the koala genome. *Nat. Genet*. **2018**, 50, 1102-1111.
12. Bolger, A.M.; Lohse, M. Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, 30, 2114-2120.
13. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al.. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol*. **2011**, 29, 644-652.
14. Buchfink, B.; Xie ,C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, 12, 59-60.

15. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol*. **1990**, 215, 403-410.
16. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, 9, 357-359.
17. Marchler-Bauer, A.; Bryant, S.H. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **2004**, 32, W327-331.
18. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol*. **2013**, 30, 772-780.
19. Capella-Gutierrez, S.; Silla-Martinez, J.M.; Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, 25, 1972-1973.
20. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, 14, 587-589.
21. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol*. **2010** 59, 307-321.
22. Wan, C.; Loader, J.; Hanger, J.; Beagley, K.; Timms, P.; Polkinghorne, A. Using quantitative polymerase chain reaction to correlate *Chlamydia pecorum* infectious load with ocular, urinary and reproductive tract disease in the koala (*Phascolarctos cinereus*). *Aust. Vet. J*. **2011**, 89, 409-412.
23. Jelocnik, M.; Islam, M.M.; Madden, D.; Jenkins, C.; Branley, J.; Carver, S.; Polkinghorne, A. Development and evaluation of rapid novel isothermal amplification assays for important veterinary pathogens: *Chlamydia psittaci* and *Chlamydia pecorum*. *PeerJ* **2017** 5, e3799.
24. Shi, M.; Lin, X.D.; Vasilakis, N.; Tian, J.H.; Li, C.X.; Chen, L.J.; Eastwood, G.; Diao, X.N.; Chen, M.H.; Chen, X.; et al. Divergent viruses discovered in arthropods and vertebrates revise the evolutionary history of the *Flaviviridae* and related viruses. *J. Virol*. **2015** 90, 659-669.
25. Remnant, E.J.; Shi, M.; Buchmann, G.; Blacquiere, T.; Holmes, E.C.; Beekman, M.; Ashe, A. A diverse range of novel RNA viruses in geographically distinct honey bee populations. *J. Virol*. **2017**, 91, e00158-17.
26. Teixeira, L.; Ferreira, A.; Ashburner, M. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol*. **2008**, 6, e2.
27. Johnson, K.N. The impact of *Wolbachia* on virus infection in mosquitoes. *Viruses* **2015**, 7, 5705-5717.
28. Bian, G.; Xu, Y.; Lu, P.; Xie, Y.; Xi, Z. The endosymbiotic bacterium *Wolbachia* induces resistance to dengue virus in *Aedes aegypti*. *PLoS Pathog*. **2010**, 6, e1000833.
29. Johnson, K.N. Bacteria and antiviral immunity in insects. *Curr. Opin. Insect. Sci*. **2015**, 8, 97-103.
30. Madden, D.; Whaite, A.; Jones, E.; Belov, K.; Timms, P.; Polkinghorne, A. Koala immunology and infectious diseases: how much can the koala bear? *Dev. Comp. Immunol*. **2018**, 82, 177-185.
31. Zhou, Y.; Holland, M.J.; Makalo, P.; Joof, H.; Roberts, C.H.; Mabey, D.C.; Bailey, R.L.; Burton, M.J.; Weinstock, G.M.; Burr, S.E. The conjunctival microbiome in health and trachomatous disease: a case control study. *Genome Med*. **2014** 6, 99.
32. Butcher, R.M.R.; Sokana, O.; Jack, K.; Kalae, E.; Sui, L.; Russell, C.; Houghton, J.; Palmer, C.; Holland, M.J.; Le Mesurier, R.T.; et al. Active trachoma cases in the Solomon Islands have varied polymicrobial community structures but do not associate with individual non-chlamydial pathogens of the eye. *Front. Med*. **2017**, 4, 251.