Title: Butcherbird polyomavirus isolated from a grey butcherbird (*Cracticus torquatus*) in Queensland, Australia.

Author: Mark D. Bennett Amber Gillett

PII: S0378-1135(13)00549-X
DOI: http://dx.doi.org/doi:10.1016/j.vetmic.2013.11.026
Reference: VETMIC 6420

To appear in: *VETMIC*

Received date: 9-9-2013
Revised date: 11-11-2013
Accepted date: 13-11-2013

Please cite this article as: Bennett, M.D., Gillett, A.,Butcherbird polyomavirus isolated from a grey butcherbird (*Cracticus torquatus*) in Queensland, Australia., *Veterinary Microbiology* (2013), http://dx.doi.org/10.1016/j.vetmic.2013.11.026

Butcherbird polyomavirus isolated from a grey butcherbird (*Cracticus torquatus*) in Queensland, Australia.

Mark D. Bennett[1], Amber Gillett[2]

1. School of Veterinary and Life Sciences, Murdoch University, South Street, Murdoch, Western Australia, Australia, 6150.
2. Australia Zoo Wildlife Hospital, Beerwah, Queensland, Australia, 4519.

Corresponding author:

Mark D. Bennett

Phone: +61 8 9360 7418

Fax: +61 8 9310 4144

Email: m.bennett@murdoch.edu.au

RH: Butcherbird polyomavirus

Abstract

A novel avian polyomavirus was detected in peri-ocular skin lesions collected from a grey butcherbird (*Cracticus torquatus*), using a combination of multiply primed rolling circle amplification, nested PCR and long range PCR. The sequence of Butcherbird polyomavirus was determined by combining next generation sequencing and primer walking techniques. The circular double-stranded DNA genome of Butcherbird polyomavirus consisted of 5084 bp, and encoded 6 open reading frames (*ORF-X*, *VP2*, *VP3*, *VP1*, *small T-antigen* and *large T-antigen*). Phylogenetic analysis placed it amongst other members of the genus *Avipolyomavirus*, most closely related to Crow polyomavirus. Next generation sequencing enabled the detection of DNA fragments similar to, but distinct from, Canarypox virus within the same lesion from which Butcherbird polyomavirus was amplified, thus confirming an avipolyomavirus-avipoxvirus co-infection in the peri-ocular skin lesions of this grey butcherbird.

Key words: *Cracticus torquatus*; polyomavirus; poxvirus; co-infection; butcherbird

Introduction

Members of the family *Polyomaviridae* are non-enveloped, icosahedral viruses measuring 40-45 nm in diameter (Imperiale & Major, 2007). Their capsids contain a single circular molecule of double stranded DNA of approximately 5000 bp, in which they encode structural proteins (VP1, VP2, VP3, and sometimes others) on one strand and transforming proteins (large, small and occasionally other T-antigens) on the opposite strand (Imperiale & Major,

2007; Johne et al., 2011). Between the start codons of these two genome regions lies a non-coding control region (NCCR), which contains the origin of DNA replication and regulatory sequences (Imperiale & Major, 2007; Johne et al., 2011).

Over the last few years, there has been a significant increase in the number of known polyomavirus (PyV) isolates from a wide range of vertebrate hosts including birds (Johne et al., 2006; Johne & Müller, 2007; Halami et al., 2010), bats (Misra et al., 2009; Fagrouch et al., 2012; Tao et al., 2013), humans (Allander et al., 2007; Gaynor et al., 2007; Feng et al., 2008; van der Meijden et al., 2010; Scuda et al., 2011; Buck et al., 2012; Lim et al., 2012; Siebrasse et al., 2012; Yu et al., 2012; Korup et al., 2013), non-human primates (Johne et al., 2005; Verschoor et al., 2008; Deuzing et al., 2010; Groenewoud et al., 2010; Leendertz et al., 2011), carnivores (Wellehan et al., 2011; Duncan et al., 2013), horses (Renshaw et al., 2012), rodents (Orba et al., 2011), and a dolphin (Anthony et al., 2013). Three genera have now been erected within the family: *Orthopolyomavirus* and *Wukipolyomavirus* for the known PyVs of mammals, and *Avipolyomavirus* for the known PyVs of birds (Johne et al., 2011).

Butcherbirds are mid-sized, black-grey and white, passerine birds classified within the genus *Cracticus* (Gillies, 1984). The genus includes 7 species, all of which are native to Australasia (Kearns et al., 2013). The grey butcherbird, *Cracticus torquatus*, is an inhabitant of every State and Territory of Australia (Kearns et al., 2013). It is regarded as common, and listed by the IUCN as a species of least concern (http://www.iucnredlist.org).

This manuscript reports the complete genetic characterisation of Butcherbird polyomavirus, a novel member of the genus *Avipolyomavirus*, in the context of co-infection with an unknown *Avipoxvirus* species.

Materials and Methods

Clinical sample

A wild adult grey butcherbird (*Cracticus torquatus*) of undetermined sex was found at Morayfield, Queensland, Australia (27.13°S; 152.91°E) on 12th July 2009, and taken to the Australia Zoo Wildlife Hospital for treatment of bilateral peri-ocular proliferative cutaneous lesions. Veterinary treatment of this patient (#19840) included antibiotic therapy (enrofloxacin 10 mg/kg intramuscularly, every 12 hours), analgesia (meloxicam 0.5 mg/kg intramuscularly, every 12 hours) and surgical removal under general anaesthesia of the proliferative "pox" lesions present around both eyes. Part of the biopsied lesional material was stored frozen (-20 °C) for molecular tests.

DNA extraction and nested PCR for polyomaviruses

Total DNA was extracted from 25 mg of finely minced lesional periocular tissue, using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's protocol, and eluted with 200 µL of elution buffer. A nested PCR protocol targeting the *VP1* open reading frame (ORF) of PyVs was used, as previously described (Johne et al., 2005). A 7 µL aliquot of the reaction mixture was mixed with loading buffer and electrophoresed through a 1% agarose gel laced with SYBR® Safe (Invitrogen) and viewed on a Gel Doc™ EZ Imager (BioRad). A PCR Clean Up kit (Qiagen) was used as per the manufacturer's instructions to purify the PCR product.

Due to a suspected low concentration of template DNA in the original (~200 µL) DNA extract, 150 µL of this material was concentrated using a PCR Clean Up kit (Qiagen), and re-eluted in 30 µL of elution buffer.

Sanger sequencing

Purified amplicon DNA was used as the template in 2 separate dideoxynucleotide chain termination sequencing reactions (Big Dye version 3.1, Applied Biosystems) with the forward and reverse primers from the second round of the nested PCR reaction described above (Johne et al., 2005). The sequence was determined using an ABI Prism Applied Biosystems 377 automatic DNA sequencer (Applied Biosystems) at the State Agriculture and Biotechnology Centre, Perth, Western Australia. Chromatogram sequencing files were edited using Chromas Lite version 2.1 (Technelysium Pty. Ltd.) and the sequences aligned manually.

Multiply primed rolling circle amplification and restriction enzyme digestion

Multiply primed rolling circle amplification (RCA) was performed using the TempliPhi™ 100 Amplification Kit (Amersham Biosciences) on a 1 µL aliquot of the concentrated sample DNA using a modified manufacturer's protocol (Rector et al., 2004). Briefly, sample template was denatured in 5 µL sample buffer at 95 °C for 3 minutes, then put on ice. To this, 5 µL of a master mix consisting of 5 µL of reaction buffer, 0.2 µL of dNTPs (25 mmol/L) and 0.2 µL of polymerase enzyme were added and mixed thoroughly by pipetting. The mixture was then incubated for approximately 60 hours at 30 °C.

Three 2 µl aliquots of the RCA product were digested with 10 units of *Bam*HI, *Eco*RI, or *Kpn*I for 5 hours at 37 °C. Following digestion, 8 µl of each digest was subjected to electrophoresis through a 1% agarose gel laced with SYBR® Safe (Invitrogen). The resulting bands were visualized on a Gel Doc™ EZ Imager (BioRad).

Long range PCR

Long range PCR was performed using a long range PCR kit (Qiagen), following the manufacturer's recommendations for amplifying targets up to 10 kbp. The experiment was performed at 3 different annealing temperatures, 50 °C, 54 °C and 58 °C, either with or without Q-Solution, for a total of 6 reactions. The primer sequences, designed using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0), were 5'-GCCTATGGAAGGTGATGACT-3', and 5'-TCCCCGGTACCTCTGTTTGC-3'. Template DNA was a mixture of 7 µL of the concentrated DNA extracted from the clinical sample, and 5 µL of the RCA product distributed evenly between the 6 reactions. Each reaction mixture was placed in a Mastercycler gradient thermocycler (Eppendorf) with the following cycling conditions: 93 °C for 3 min; then 40 cycles of 93 °C for 15 s, annealing for 30 s, 68 °C for 8 min; then 14 °C hold. A 7 µL aliquot of the reaction mixture was electrophoresed through a 1% agarose gel laced with SYBR® Safe (Invitrogen) and viewed on a Gel Doc™ EZ Imager (BioRad). A PCR Clean Up kit (Qiagen) was used as per the manufacturer's instructions to purify the Long Range PCR product.

Primer walking

The long range PCR primers were used in the first round of Sanger sequencing. Subsequently determined sequence information was used to design more primers, with the assistance of the online Primer3 tool (http://bioinfo.ut.ee/primer3-0.4.0). The primers used were: GBF2 CAACACAAAGCATCTGTTGT; GBR2 CCTGTGAATAATAGTCTTTGA; GBF3 TGCTTTCTTTGCTTTTCCAT; and GBR3 CCAGTATCTGCAGACTCCTT.

Next generation sequencing

The DNA library was prepared using the Roche Rapid Library kit according to the Rapid Library Preparation Method Manual. Purified long range PCR amplified DNA was fragmented to the appropriate size range using nebulization. Universal Lib L sequencing adapters were annealed to the fragment ends and fragments smaller than 300 bp were removed with Ampure magnetic beads. The fragment size distribution of the DNA library was checked using an Agilent 2100 Bioanalyzer High Sensitivity DNA chip and the library was quantified using a DTX 880 fluorescent plate reader. Emulsion PCR was carried out according to the emPCR Amplification Method Manual - Lib-L using the GSJR titanium emPCR kit Lib L. Each bead was mixed with 1.2 DNA molecules in the emulsion PCR, resulting in 20% enrichment. Sequencing was carried out on a GS Junior sequencer using GSJR titanium sequencing kit and GSJR Pico titre plate according to the Roche Sequencing Method Manual.

Sequence analysis and archiving

Next generation sequence (NGS) data were first assembled using the GS De Novo Assembler Version 2.8 (Roche, 454 Life Sciences). Settings used included minimum read length: 20; seed length: 16; seed count: 1; minimum overlap length: 40, minimum overlap identity: 90%, alignment identity score: 2; and alignment difference score: -3. Subsequently, GS Reference Mapper Version 2.8 (Roche, 454 Life Sciences) was used to map the raw NGS data against Canarypox virus strain ATCC VR-111 (AY318871), which allowed bioinformatic subtraction of reads mapping to the avipoxvirus present within the sample, followed by mapping of unmapped reads against a draft Butcherbird PyV sequence obtained through primer walking. Settings for the GS Reference Mapper software were identical to those for the GS De Novo Assembler software with the additional items of hits per seed limit: 70; and repeat score threshold: 12. Sequence information obtained from next generation, primer walking and initial Sanger sequencing was imported into BioEdit version 7.0.9.0 and aligned manually. Open reading frames were detected using the NCBI online ORF Finder application (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) combined with manual inspection of the sequence. Splice donor and acceptor sites were predicted using the online software NNSPLICE version 0.9 (http://www.fruitfly.org/seq_tools/splice.html). The online Basic Local Alignment Search Tool (BLAST) was used to interrogate GenBank for similar previously archived sequences. Identified ORF nucleotide sequences were translated using BioEdit version 7.0.9.0, and the predicted proteins' theoretical isoelectric points and molecular masses were estimated using the online Compute pI/Mw tool (http://www.expasy.ch/tools). Amino acid sequences of coding regions and nucleotide sequences of non-coding regions were searched manually for established motifs. The Butcherbird polyomavirus sequence was annotated and submitted to GenBank using the online BankIt tool.

Phylogenetic analysis

The predicted amino acid sequences of VP2, VP3, VP1, small and large T-antigens from the current isolate and 47 previously established PyVs of avians and mammals, and the T-antigen encoding regions of the bandicoot papillomatosis carcinomatosis viruses (BPCVs) types 1 and 2 (Table 1) were imported into BioEdit version 7.0.9.0 and aligned using ClustalW. The amino acid alignment was manually edited to remove ambiguity and to produce two separate sequences, 1 including the late protein-encoding regions and 1 including the T-antigen encoding regions. The resulting amino acid sequences were de-translated back to the original nucleotide sequences, and these were imported into MEGA version 5.2.2 for phylogenetic analysis. Initial trees were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach. A Maximum Likelihood consensus tree was subsequently inferred using the General Time Reversible model and 1000 bootstrap replicates.

Pairwise analysis

The appropriate amino acid sequences and their related nucleotide sequences from the VP2/3, VP1, small T-antigen and large T-antigen from Butcherbird PyV, 5 avian PyVs, 11 selected mammalian PyVs and the small and large T-antigens of the 2 bandicoot papillomatosis carcinomatosis viruses were imported into BioEdit version 7.0.9.0 and aligned unsupervised by ClustalW at the amino acid level and at the nucleotide level. The percentage identity between Butcherbird PyV and the other 16-18 isolates was calculated by the BioEdit software.

Results

Genome amplification and sequencing

The nested PCR for amplifying a wide range of PyVs yielded a single bright band of approximately 230 bp, which Sanger sequencing revealed to be similar, but not identical, to Crow PyV.

Restriction enzyme digestion of the RCA product with *Bam*HI produced no bands, whereas with *Eco*RI and *Kpn*I, several faint bands were visible (Fig. 1). Although the results were not clear enough to accurately estimate the sizes of the restriction digest fragments, they indicated that a small circular DNA genome was present in the DNA sample.

A single bright band of approximately 5 kbp was produced without Q solution, and using an annealing temperature of 54 °C (Fig. 1). The 5 other reaction conditions tested failed. The purified product of this experiment was used for both primer walking and next generation sequencing experiments.

Next generation sequencing data were initially sorted using de novo sequencing. Surprisingly, this approach yielded several hundred contigs, none of which were related to PyVs. Rather, the overwhelming majority of contigs were similar to Canarypox virus strain ATCC VR-111 (AY318871). A random selection of 20 contigs had an average sequence identity of 95.5% with Canarypox virus strain ATCC VR-111, with a range of 84-99%. The next generation sequencing data were then imported into GS reference mapper version 2.8, and mapped against the Canarypox virus sequence. Mapped sequences were bioinformatically subtracted from the data set, and the remainder then mapped against the

Butcherbird PyV draft sequence obtained from primer walking results. Following this procedure, 99.3% of unmapped sequences were successfully mapped to the Butcherbird PyV draft sequence, at an average coverage depth of 9,338.23 reads and a maximum coverage depth of 23,782.

Sequence analysis and archiving

The genome of Butcherbird PyV (GenBank accession number KF360862) was a single circular molecule of double-stranded DNA that consisted of 5084 bp and had a G+C content of 45.7%. The first nucleotide in the genome sequence was set as the 1 immediately after the start codon for the T-antigens on the complementary DNA strand. Open reading frames typical of avian PyVs were detected, including an *ORF-X*, *VP2*, *VP3*, *VP1*, *small T-antigen* and *large T-antigen* (Fig. 2).

Upstream of *VP2*, between nucleotides 281 and 765 was a region consistent with the *ORF-X*s of avian PyVs. An intron was predicted from nucleotide 393 to 463 inclusive. The predicted product of this *ORF-X* was estimated to have 137 amino acid residues, high proline content (21/137; 15.22%), an approximate molecular mass of 14.0 kDa and an isoelectric point at 4.57. When the nucleotide sequence was analysed, 78% identity was detected with part of the Crow PyV whole genome sequence, and when the amino acid sequence was analysed, there was 52% identity with Goose hemorrhagic PyV's ORF-X and 36% identity with Crow PyV's ORF-X.

The *VP2* ORF was found between nucleotides 929 and 1924, and was predicted to encode a 331 amino acid long VP2 protein with an isoelectric point at 5.44 and molecular mass of 36.0

kDa. Contained within the 3' end of the *VP2* ORF (1262-1924) was a region predicted to encode VP3, a 220 amino acid long 24.9 kDa protein with an isoelectric point at 9.44.

Overlapping the very end of the *VP2/VP3* ORFs, from nucleotide position 1809 to 2882, there was an ORF predicted to encode VP1. This 357 amino acid long major capsid protein had a predicted isoelectric point at 6.35 and molecular mass of 38.2 kDa.

On the opposite genomic strand, from nucleotide position 5084 to 4572, was an ORF predicted to encode a small T-antigen. Butcherbird PyV small T-antigen consisted of 170 amino acids, contained the E/D-xxx-L-x-E/D-L-xx-L/I motif (^{5}EIRELRELLGL^{15}), the conserved hexapeptide DnaJ motif ^{36}HPDKGG^{41}, a retinoblastoma protein binding motif (^{70}LFCDE^{74}), and had a predicted molecular mass of 19.2 kDa and isoelectric point at 5.09.

The predicted large T-antigen-encoding ORF of Butcherbird PyV was found between nucleotide positions 5084 and 2979. It overlapped the start of the small T-antigen region, but an intron was predicted between nucleotide positions 4754 and 4572. The predicted Butcherbird PyV large T-antigen consisted of 640 amino acid residues, with an approximate molecular mass of 71.4 kDa and an isoelectric point at 6.16. It included the motifs mentioned for the small T-antigen above, as well as a putative zinc-finger motif starting at amino acid residue 303, putative ATPase binding domain motif ^{436}GAVNTGKT^{443}, and the highly conserved G-xxx-VNLE motif, ^{515}GAVPVNLE^{522}.

Between the 5' ends of the T-antigen ORFs and *ORF-X*, there was a 280 nucleotide non-coding control region (NCCR) which included a putative avian PyV T-antigen binding site, CC(A/T)$_6$GG (Luo et al., 1994). There was a short (95 nucleotide) non-coding region between the 3' ends of the *VP1* ORF and the *large T-antigen* ORF, and a 163 nucleotide non-coding region between the end of *ORF-X* and the start of *VP2*.

Phylogenetic analysis

Forty-eight PyV sequences were selected for inclusion in the phylogenetic analysis (Table 1). Following alignment and manual editing to remove data gaps and ambiguous areas of alignment, a concatenated VP2/3/1 nucleotide sequence consisting of 1026 sites was constructed that corresponded to nucleotide positions 929-952, 959-979, 1037-1060, 1079-1138, 1304-1342, 1355-1390, 1508-1522, 1667-1699, 1706-1765, 1871-1921, 1923-1973, 2025-2168, 2205-2312, 2346-2363, 2373-2597, 2619-2720 and 2766-2780 relative to the Butcherbird PyV sequence. Similarly, a concatenated T-antigens nucleotide sequence consisting of 966 sites was constructed, corresponding to positions 3120-3183, 3190-3413, 3441-3627, 3745-3914, 4002-4095, 4189-4297 and 4353-4469 of the Butcherbird PyV sequence.

Butcherbird PyV consistently clustered with the other known avian PyVs, and was most closely related to Crow PyV (Fig. 3, 4). Based on a whole-genome pairwise alignment between Crow PyV and the current isolate, a sequence identity of 63.2% was calculated. Analysis of the T-antigens-encoding part of the genomes confirmed members of *Avipolyomavirus* formed a separate clade from the polyomaviruses of mammals (Fig. 3). The most closely related T-antigen sequences obtained from mammalian hosts came from the bandicoot papillomatosis carcinomatosis viruses (BPCVs), which are not currently classified within *Polyomaviridae* (Fig. 3).

The phylogenetic tree based on the structural proteins-encoding part of the genomes unexpectedly showed Equine PyV amongst the cluster of *Avipolyomavirus* isolates. There

was low bootstrap support at the basal part of this tree, suggesting this region of the genome may be unreliable for inferring deep relationships between isolates (Fig. 4).

Pairwise alignments

Butcherbird PyV predicted proteins were generally most similar to other known avian PyVs, both at the amino acid and their coding nucleotide sequence levels (Table 2). This observation was most clearly demonstrated when the results for the T-antigens were examined, whereas there were several instances of mammalian PyV sequences sharing greater sequence similarity with Butcherbird PyV than other avian PyVs when looking at their predicted structural proteins. The large T-antigens of the BPCVs, which are natural hybrid viruses encoding papillomavirus-like structural proteins and PyV-like non-structural proteins, clearly occupied the middle ground between avian and eutherian PyVs at both the amino acid and nucleotide sequence levels, but this was not obvious if the analysis was restricted to include only the small T-antigen region.

Discussion

The Butcherbird PyV genome organisation was typical of members of the genus *Avipolyomavirus*. It had an *ORF-X* gene with a predicted splice site, which encoded a proline-rich product between the NCCR and late ORFs; an avian PyV large T-antigen binding site in the NCCR, $CC(A/T)_6GG$, rather than the mammalian motif GAGGC; and lacked both the N-terminal nuclear localization signal within VP1, and CxCxxC PP2A binding site within the small T-antigen (Halami et al., 2010). *In silico* prediction tools were employed to identify the

likely splice donor and acceptor sites in the Butcherbird PyV genes *ORF-X* and *large T-antigen*. While the predicted sites scored highly according to the computer algorithm, and the predicted proteins generated by such splices were consistent with other known PyV *ORF-X* product and large T-antigen proteins, these results should be confirmed by sequencing of actual virus encoded mRNA species.

Butcherbird PyV is clearly most closely related to Crow PyV, according to the phylogenetic and pairwise analyses conducted. It is tempting to hypothesize that the relatedness of these two avian PyVs is a reflection of the evolutionary histories of their hosts. Butcherbirds and crows are both members of superfamily Corvoidea (http://sn2000.taxonomy.nl/), and thus, it is possible that this represents a case of virus-host co-speciation. However, examination of the phylogenetics of the mammalian PyVs showed that virus-host co-speciation alone could not explain the evolutionary relationships between all PyV isolates (Krumbholz et al., 2009; Warden & Lacey, 2012; Tao et al., 2013). Discovery and characterisation of more avian PyV isolates will help to resolve the question of the evolutionary mechanisms of greatest importance within the genus *Avipolyomavirus*.

Long held ideas regarding the involvement of PyVs in manifestations of animal disease are now being challenged, as new members of the family are discovered in a wider range of animal hosts and in unexpected anatomical compartments. Previously, it was owned that mammalian PyV infections were generally persistent and clinically silent, so long as the host's immune system was properly functional, while PyV infections of birds typically led to severe, acute and often fatal disease (Johne et al., 2006; Johne et al., 2011). Currently available evidence provides examples of a PyV infection in association with acute ulcerative respiratory tract disease in mammals, for instance, the common dolphin (*Delphinus delphis*) (Anthony et al., 2013). The current isolate, Butcherbird PyV, also provides evidence against

the current paradigm. The genome was discovered serendipitously in an individual that subsequently recovered well enough to be released to the wild.

It is striking that the number of PyV species attributed to humans far outweighs those for any other host species. This is probably a reflection of the magnitude of the resources employed in novel virus discovery in humans compared with any other animal host species. Based on these observations, it appears reasonable to predict a plethora of novel PyV isolates in nature, with each vertebrate host species likely harbouring its own PyV milieu. On the assumption that our current understanding of PyV-host interactions is informed by only a mere fraction of the total diversity available for study, it seems unwise to cling fastidiously to the notions of the past.

In this case, the obvious peri-ocular skin lesions were clinically consistent with an avipoxvirus infection, and therefore the discovery of Butcherbird PyV may have been an entirely serendipitous event, unrelated to the poxvirus infection itself. On the other hand, the poxvirus infection may have led to immunosuppression of the butcherbird, such that a formerly latent infection became productive; or else, a pre-existing general state of immune suppression in this individual butcherbird might have facilitated both elements of this co-infection. Shivaprasad et al. (2009) reported on adult canaries (*Serinus canaria*) with concurrent canarypox virus – PyV infections, and speculated that the PyV infections may have exacerbated clinical signs and mortality rates in affected individuals. Halami et al. (2010), subsequently discovered canary PyV, which was associated with fatal disease in young canaries. Whether the canary PyV of Halami *et al*. (2010) is the same as that mentioned by Shivaprasad *et al*. (2009) is unknown. There is currently no compelling evidence that Butcherbird PyV is associated with any particular pathology in butcherbirds, however further studies are required to more fully elucidate this point.

Butcherbird PyV was confirmed as a separate species according to the criteria established by the *Polyomaviridae* study group of the International Committee on Taxonomy of Viruses (ICTV) (Johne et al., 2011). It shares only 63.2% identity with its closest known relative, on a whole genome basis, and host species and phylogenetic evidence both support its inclusion in the genus *Avipolyomavirus* as a new and distinct species.

Acknowledgments

This study was funded by the Murdoch University School of Veterinary and Life Sciences. The authors would like to acknowledge Claude Lacasse for her assistance with collecting the sample for analysis, and Frances Brigg and David Berryman for assistance with sequencing and bioinformatics analysis.

References

Allander, T., Andreasson, K., Gupta, S., Bjerkner, A., Bogdanovic, G., Persson, M.A.A., Dalianis, T., Ramquist, T., Andersson, B., 2007. Identification of a third human polyomavirus. J. Virol. 81, 4130–4136.

Anthony, S.J., St. Leger, J.A., Navarrete-Macias, I., Nilson, E., Sanchez-Leon, M., Seimon, T., Jain, K., Karesh, W., Daszak, P., Briese, T., Lipkin, W.I., 2013. Identification of a novel

cetacean polyomavirus from a common dolphin (*Delphinus delphis*) with tracheobronchitis. PLOS ONE 8, e68239.

Buck, C.B., Phan, G.Q., Raiji, M.T., Murphy, P.M., McDermott, D.H., McBride, A.A., 2012. Complete genome sequence of a tenth human polyomavirus. J. Virol. 86, 10887.

Deuzing, I., Fagrouch, Z., Groenewoud, M.J., Niphuis, H., Kondova, I., Bogers, W., Verschoor, E.J., 2010. Detection and characterization of two chimpanzee polyomavirus genotypes from different subspecies. Virol. J. 7, 347-354.

Duncan, C., Goldstein, T., Hearne, C., Gelatt, T., Spraker, T., 2013. Novel polyomavirus infection in the placenta of a northern fur seal (*Callorhinus ursinus*) on the Pribilof Islands, Alaska, USA. J. Wildl. Dis. 49, 163-167.

Fagrouch, Z., Sarwari, R., Lavergne, A., Delaval, M., de Thoisy, B., Lacoste, V., Verschoor, E.J., 2012. Novel polyomaviruses in South American bats and their relationship to other members of the family *Polyomaviridae*. J. Gen. Virol. 93, 2652-2657.

Feng, H., Shuda, M., Chang, Y., Moore, P.S., 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 319, 1096–1100.

Gaynor, A.M., Nissen, M.D., Whiley, D.M., Mackay, I.M., Lambert, S.B., Wu, G., Brennan, D.C., Storch, G.A., Sloots, T.P., Wang, D., 2007. Identification of a novel polyomavirus from patients with acute respiratory tract infections. PLOS Pathog. 3, e64.

Gillies, B., 1984. Cracticidae. In: Simpson, K., Day, N., (Eds.), The birds of Australia: a book of identification., pp. 266-267.

Groenewoud, M.J., Fagrouch, Z., van Gessel, S., Niphuis, H., Bulavaite, A., Warren, K.S., Heeney, J.L., Verschoor, E.J., 2010. Characterization of novel polyomaviruses from Bornean and Sumatran orang-utans. J. Gen. Virol. 91, 653–658.

Halami, M.Y., Dorrestein, G.M., Couteel, P., Heckel, G., Müller, H., Johne, R., 2010. Whole genome characterization of a novel polyomavirus detected in fatally diseased canary birds. J. Gen. Virol. 91, 3016–3022.

Imperiale, M.J., Major, E.O., 2007. *Polyomaviridae*. In: Knipe, D.M., Howley, P.M. (Eds.), Field's Virology, 5th ed., vol 2. Lippincott Williams and Wilkins, Philadelphia, PA, pp. 2263-2298.

Johne, R., Buck, C.B., Allander, T., Atwood, W.J., Garcea, R.L., Imperiale, M.J., Major, E.O., Ramqvist, T., Norkin, L.C., 2011. Taxonomical developments in the family *Polyomaviridae*. Arch. Virol. 156, 1627–1634.

Johne, R., Enderlein, D., Nieper, H., Müller, H., 2005. Novel polyomavirus detected in the feces of a chimpanzee by nested broadspectrum PCR. J. Virol. 79, 3883–3887.

Johne, R., Müller, H., 2007. Polyomaviruses of birds: etiologic agents of inflammatory diseases in a tumor virus family. J. Virol. 81, 11554–11559.

Johne, R., Wittig, W., Fernandez-de-Luco, D., Höfle, U., Müller, H., 2006. Characterization of two novel polyomaviruses of birds by using multiply primed rolling-circle amplification of their genomes. J. Virol. 80, 3523–3531.

Kearns, A.M., Joseph, L., Cook, L.G., 2013. A multilocus coalescent analysis of the speciational history of Australo-Papuan butcherbirds and their allies. Mol. Phylogenet. Evol. 66, 941-952.

Korup, S., Rietscher, J., Calvignac-Spencer, S., Trusch, F., Hofmann, J., Moens, U., Sauer, I., Voigt, S., Schmuck, R., Ehlers, B., 2013. Identification of a novel human polyomavirus in organs of the gastrointestinal tract. PLOS ONE 8, e58021

Krumbholz, A., Bininda-Emonds, O.R.P., Wutzler, P., Zell, R., 2009. Phylogenetics, evolution, and medical importance of polyomaviruses. Infect. Genet. Evol. 9, 784-799.

Leendertz, F.H., Scuda, N., Cameron, K.N., Kidega, T., Zuberbühler, K., Leendertz, S.A.J., Couacy-Hymann, E., Boesch, C., Calvignac, S., Ehlers, B., 2011. African great apes are naturally infected with polyomaviruses closely related to Merkel cell polyomavirus. J. Virol. 85, 916–924.

Lim, E.S., Reyes, A., Antonio, M., Saha, D., Ikumapayi, U.N., Adeyemi, M., Stine, O.C., Skelton, R., Brennan, D.C., Mkakosya, R.S., Manary, M.J., Gordon, J.I., Wang, D., 2012. Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. Virology 436, 295-303.

Luo, D., Müller, H., Tang, X. B.,Hobom, H., 1994. Expression and DNA binding of budgerigar fledgling disease virus large T antigen. J. Gen. Virol. 75, 1267-1280.

Misra, V., Dumonceaux, T., Dubois, J., Willis, C., Nadin-Davis, S., Severini, A., Wandeler, A., Lindsay, R., Artsob, H., 2009. Detection of polyoma and corona viruses in bats of Canada. J. Gen. Virol. 90, 2015–2022.

Orba, Y., Kobayashi, S., Nakamura, I., Ishii, A., Hang'ombe, B.M., Mweene, A.S., Thomas, Y., Kimura, T., Sawa, H., 2011. Detection and characterization of a novel polyomavirus in wild rodents. J. Gen. Virol. 92, 789-795.

Rector, A., Tachezy, R., Van Ranst, M., 2004. A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. J. Virol. 78, 4993-4998.

Renshaw, R.W., Wise, A.G., Maes, R.K., Dubovi, E.J., 2012. Complete genome sequence of a polyomavirus isolated from horses. J. Virol. 86, 8903.

Scuda, N., Hofmann, J., Calvignac-Spencer, S., Ruprecht, K., Liman, P., Kühn, J., Hengel, H., Ehlers, B., 2011. A novel human polyomavirus closely related to the african green monkey derived lymphotropic polyomavirus. J. Virol. 85, 4586–4590.

Shivaprasad, H.L., Kim, T., Tripathy, D., Woolcock, P.R., Uzal, F., 2009. Unusual pathology of canary poxvirus infection associated with high mortality in young and adult breeder canaries (*Serinus canaria*). Avian Pathol. 38, 311-316.

Siebrasse, E.A., Reyes, A., Lim, E.S., Zhao, G., Mkakosya, R.S., Manary, M.J., Gordon, J.I., Wang, D., 2012. Identification of MW polyomavirus, a novel polyomavirus in human stool. J. Virol. 86, 10321–10326.

Tao, Y., Shi, M., Conrardy, C., Kuzmin, I.V., Recuenco, S., Agwanda, B., Alvarez, D.A., Ellison, J.A., Gilbert, A.T., Moran, D., Niezgoda, M., Lindblade, K.A., Holmes, E.C., Breiman, R.F., Rupprecht, C.E., Tong, S., 2013. Discovery of diverse polyomaviruses in bats and the evolutionary history of the *Polyomaviridae*. J. Gen. Virol. 94, 738-748.

van der Meijden, E., Janssens, R.W., Lauber, C., Bouwes Bavinck, J.N., Gorbalenya, A.E., Feltkamp, M.C.W., 2010. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. PLOS Pathog. 6, e1001024.

Verschoor, E.J., Groenewoud, M.J., Fagrouch, Z., Kewalapat, A., van Gessel, S., Kik, M.J., Heeney, J.L., 2008. Molecular characterization of the first polyomavirus from a New World primate: squirrel monkey polyomavirus. J. Gen. Virol. 89, 130–137.

Warden, C.D., Lacey, S.F., 2012. Updated phylogenetic analysis of polyomavirus-host co-evolution. J. Bioinform. Res. 1, 46-49.

Wellehan, J.F. Jr., Rivera, R., Archer, L.L., Benham, C., Muller, J.K., Colegrove, K.M., Gulland, F.M., St. Leger, J.A., Venn-Watson, S.K., Nollens, H.H., 2011. Characterization of California sea lion polyomavirus 1: expansion of the known host range of the *Polyomaviridae* to carnivora. Infect. Genet. Evol. 11, 987–996.

Yu, G., Greninger, A.L., Isa, P., Phan, T.G., Martinez, M.A., de la Luz Sanchez, M., Contreras, J.F., Santos-Preciado, J.I., Parsonnet, J., Miller, S., DeRisi, J.L., Delwart, E., Arias, C.F., Chiu, C.Y., 2012. Discovery of a novel polyomavirus in acute diarrheal samples from children. PLOS ONE 7, e49449.

Table 1. List of the 48 polyomavirus species and 2 bandicoot papillomatosis carcinomatosis viruses included in phylogenetic and pairwise analyses, and their GenBank accession numbers.

Table 2. Pairwise analysis comparing selected coding regions of the Butcherbird PyV with the relevant parts of 16 selected PyVs and the 2 known bandicoot papillomatosis carcinomatosis viruses.

Figure captions

Figure 1. Gel pictures showing amplification of Butcherbird polyomavirus. **A:** Multiply primed rolling circle amplification products digested with *Kpn*I (lane 2), *Eco*RI (lane 3) and

*Bam*HI (lane 4). A 1 kb DNA ladder marker (Axygen 300-10,000) is in lanes 1 & 11, negative water controls digested with *Kpn*I, *Eco*RI and *Bam*HI in lanes 5-7, and positive plasmid controls digested with *Kpn*I, *Eco*RI and *Bam*HI in lanes 8-10. **B:** Long range PCR products: 1 kb DNA ladder marker (Axygen 300-10,000) (lanes 1 & 8), water negative control (lane 2), reactions without Q-solution at 50°C, 54°C and 58°C (lanes 3-5), and reactions with Q-solution at 50°C, and 54°C (lanes 6 & 7).

Figure 2. Linearized genome map of Butcherbird polyomavirus (KF360862). The Butcherbird PyV genome consists of a 5084 bp circular molecule of double-stranded DNA. The nucleotide sequence of the non-coding control region (NCCR) contains the T-antigen binding site motif of avian PyVs (underlined). Open reading frames typical of avian PyVs are present: *ORF-X*, *VP2*, *VP3*, *VP1*, *large T-antigen* and *small T-antigen*. Predicted introns in *ORF-X* and *large T-antigen* are indicated by fine lines connecting exons displayed in bold.

Figure 3. Phylogenetic tree displaying the relationships between the aligned T-antigens-encoding regions of the genomes of 48 selected members of *Polyomaviridae* and 2 bandicoot papillomatosis carcinomatosis viruses. Only bootstrap support values of ≥75% are annotated at the nodes. The scale bar indicates the genetic distance in nucleotide substitutions per site. Butcherbird PyV is indicated by the arrow.

Figure 4. Phylogenetic tree displaying the relationships between the aligned VP2/3/1-encoding regions of the genomes of 48 selected members of *Polyomaviridae*. Only bootstrap

support values of ≥75% are annotated at the nodes. The scale bar indicates the genetic distance in nucleotide substitutions per site. Butcherbird PyV is indicated by the arrow.

Virus	Accession number	Virus	Accession number
Butcherbird PyV	KF360862	*Cebus albifrons* PyV 1	NC_019854
Crow PyV	NC_007922	*Mastomys* PyV	AB588640
Finch PyV	NC_007923	California sea lion PyV 1	NC_013796
Canary PyV	NC_017085	Bat (Myotis) PyV	NC_011310
Goose hemorrhagic PyV	NC_004800	*Pteronotus* PyV	NC_020070
Avian PyV 1	NC_004764	Equine PyV	NC_017982
Squirrel monkey PyV	NC_009951	Yellow baboon PyV 2	AB767295
BK PyV	NC_001538	Vervet monkey PyV 1	NC_019844
Cercopithecus erythrotis PyV 1	JX159985	*Piliocolobus rufomitratus* PyV 1	NC_019850
MW PyV	NC_018102	Yellow baboon PyV 1	AB767294
B-lymphotropic PyV	NC_004763	*Macaca fascicularis* PyV 1	NC_019851
Human PyV 9	NC_015150	*Miniopterus* PyV	NC_020069
Human PyV 12	NC_020890	*Ateles paniscus* PyV 1	NC_019853
Orangutan PyV	NC_013439	Orangutan PyV pi	FN356901
Pan troglodytes schweinfurthii PyV 2	NC_019858	*Pan troglodytes verus* PyV 3	NC_019855
Vervet monkey PyV 2	AB767299	STL PyV	NC_020106
Chimpanzee PyV	NC_014743	*Otomops* PyV 2	NC_020066
Chaerephon PyV 1	NC_020065	Raccoon PyV	JQ178241
Pan troglodytes verus PyV 4	NC_019856	*Eidolon* PyV 1	NC_020068
Otomops PyV 1	NC_020071	*Cardioderma* PyV	NC_020067
JC PyV	NC_001699	SV40 PyV	NC_001669
WU PyV	NC_009539	KI PyV	NC_009238
Human PyV 6	NC_014406	Human PyV 7	NC_014407
Pan troglodytes verus PyV 2a	HQ385748	*Pan troglodytes verus* PyV 5	NC_019857
BPCV1	NC_010107	BPCV2	NC_010817

1

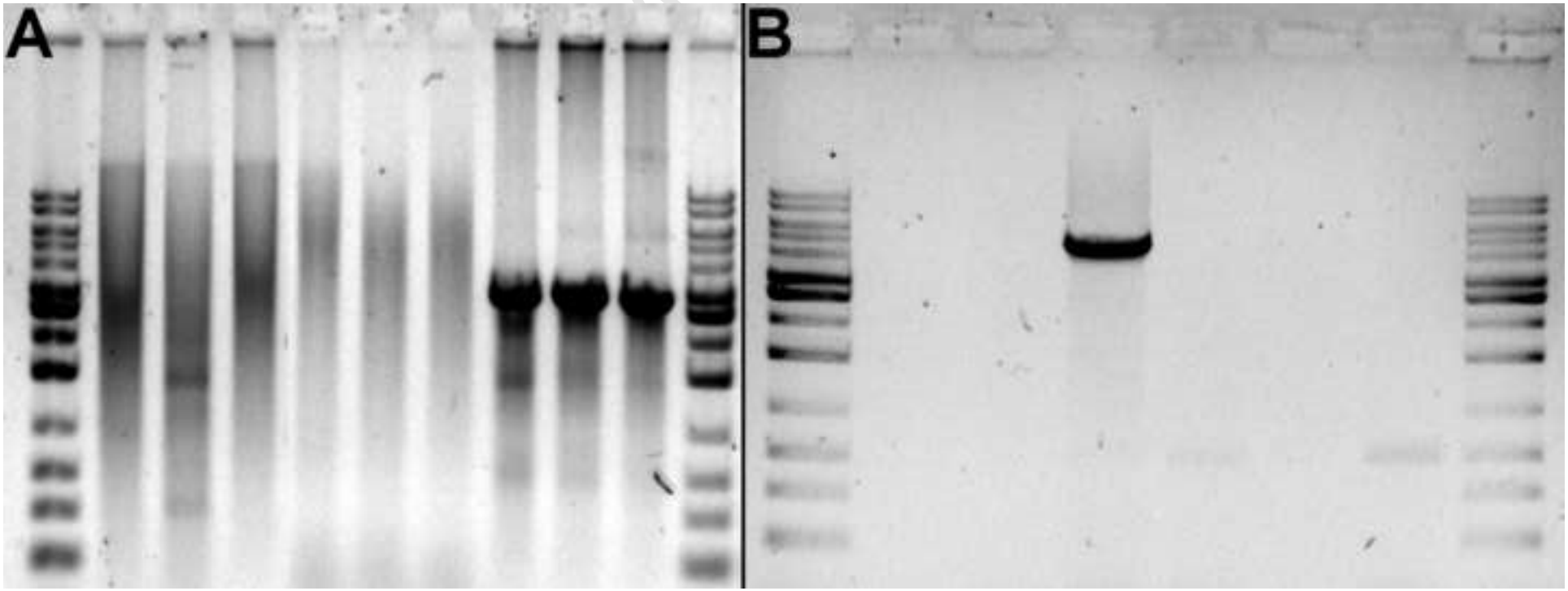
	Small T-antigen		Large T-antigen		VP1		VP2/3	
	aa	n	aa	n	aa	n	aa	n
Crow PyV	63.2	64.2	68.6	67.6	77.2	68.6	55.1	57.6
Goose hemorrhagic PyV	52.9	57.2	59.7	61.8	73.2	64.0	46.3	56.6
Finch PyV	37.2	42.6	46.2	49.7	61.6	58.2	33.2	40.7
Canary PyV	35.1	39.2	48.2	51.3	59.7	57.0	26.7	35.0
Avian PyV 1	34.5	41.0	43.8	47.5	55.0	53.8	27.0	37.3
STL PyV	15.8	30.7	24.8	32.1	41.2	44.6	24.3	36.0
Squirrel monkey PyV	16.8	30.1	24.3	32.0	52.7	55.4	27.4	43.1
Equine PyV	16.2	29.2	22.8	31.8	60.3	59.0	34.3	40.4
California sea lion PyV 1	17.3	30.5	24.0	31.3	42.5	42.0	31.1	44.6
BK PyV	19.4	31.5	22.6	30.6	51.5	55.3	32.7	39.6
Orangutan PyV	21.3	32.5	21.0	32.3	53.2	55.9	30.6	42.5
Cardioderma PyV	19.6	33.6	20.5	28.8	39.5	42.2	14.9	23.4
MW PyV	19.4	30.2	24.4	33.1	44.5	46.3	26.7	37.6
Pan troglodytes verus PyV 3	20.1	30.8	21.1	31.0	54.9	52.9	31.4	35.0
Human PyV 9	18.4	30.7	21.3	31.8	58.0	55.7	29.0	38.2
Human PyV 12	15.6	29.1	19.4	32.2	56.1	56.0	24.0	34.7
BPCV1	19.6	28.4	35.7	42.7	-	-	-	-
BPCV2	19.0	30.1	35.1	42.7	-	-	-	-

2

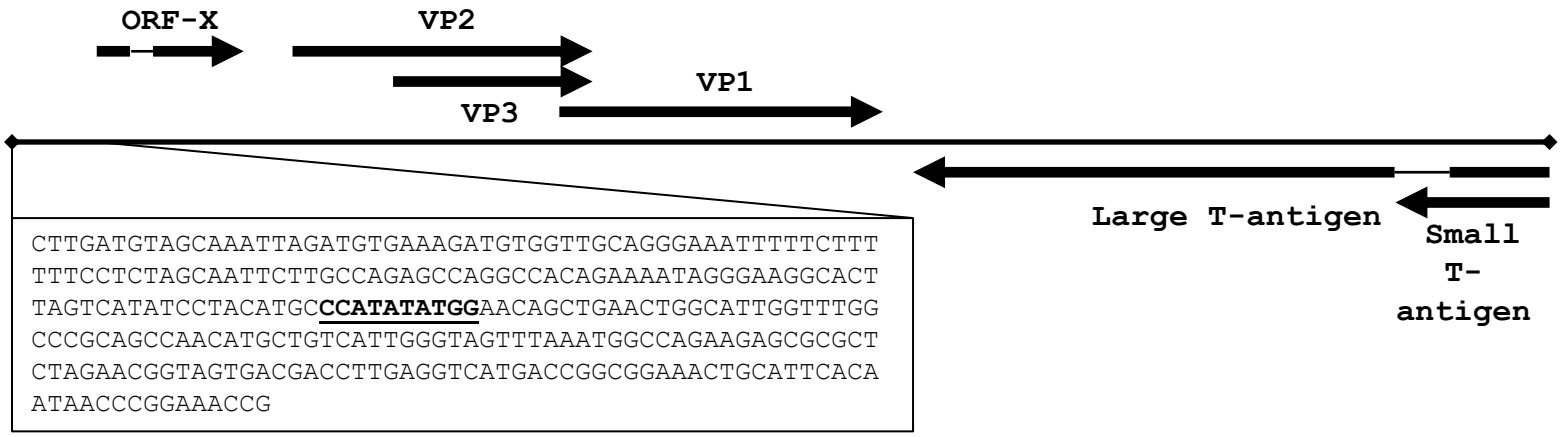
3 aa = amino acid; n = nucleotide

4

Figure



Figure


ORF-X
VP2
VP1
VP3
Large T-antigen
Small T-antigen
CTTGATGTAGCAAATTAGATGTGAAAGATGTGGTTGCAGGGAAATTTTTCTTT
TTTCCTCTAGCAATTCTTGCCAGAGCCAGGCCACAGAAAATAGGGAAGGCACT
TAGTCATATCCTACATGCCCATATATGGAACAGCTGAACTGGCATTGGTTTGG
CCCGCAGCCAACATGCTGTCATTGGGTAGTTTAAATGGCCAGAAGAGCGCGCT
CTAGAACGGTAGTGACGACCTTGAGGTCATGACCGGCGGAAACTGCATTCACA
ATAACCCGGAAACCG