Identification of an avian polyomavirus associated with Adélie penguins (Pygoscelis adeliae)

Arvind Varsani,1,2,3 Elizabeth L. Porzig,4 Scott Jennings,5 Simona Kraberger,1 Kata Farkas,1 Laurel Julian,1 Melanie Massaro,6 Grant Ballard4 and David G. Ainley7

Correspondence
Arvind Varsani
arvind.varsani@canterbury.ac.nz

1School of Biological Sciences and Biomolecular Interaction Centre, University of Canterbury, Private Bag 4800, Christchurch, 8140, New Zealand
2Department of Plant Pathology and Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA
3Electron Microscope Unit, Division of Medical Biochemistry, Department of Clinical Laboratory Sciences, University of Cape Town, Observatory, 7700, South Africa
4Point Blue Conservation Science, Petaluma, CA, USA
5Department of Fisheries and Wildlife, Oregon Cooperative Fish and Wildlife Research Unit, US Geological Survey, Oregon State University, Corvallis, OR, USA
6School of Environmental Sciences, Charles Sturt University, Albury, NSW, Australia
7HT Harvey and Associates, Los Gatos, CA, USA

Little is known about viruses associated with Antarctic animals, although they are probably widespread. We recovered a novel polyomavirus from Adélie penguin (Pygoscelis adeliae) faecal matter sampled in a subcolony at Cape Royds, Ross Island, Antarctica. The 4988 nt Adélie penguin polyomavirus (AdPyV) has a typical polyomavirus genome organization with three ORFs that encoded capsid proteins on the one strand and two non-structural protein-coding ORFs on the complementary strand. The genome of AdPyV shared ~60% pairwise identity with all avipolyomaviruses. Maximum-likelihood phylogenetic analysis of the large T-antigen (T-Ag) amino acid sequences showed that the T-Ag of AdPyV clustered with those of avipolyomaviruses, sharing between 48 and 52% identities. Only three viruses associated with Adélie penguins have been identified at a genomic level, avian influenza virus subtype H11N2 from the Antarctic Peninsula and, respectively, Pygoscelis adeliae papillomavirus and AdPyV from capes Crozier and Royds on Ross Island.

There is extremely little information about pathogens and parasites associated with Antarctic animals (Barbosa & Palacios, 2009; Kerry & Riddle, 2009). Of all Antarctic vertebrates, the penguins are possibly the best studied. Antarctic penguins have been showing signs of disease of unknown pathology in recent years, in particular unexplained incursions of feather loss in Adélie penguins (Pygoscelis adeliae) and Emperor penguins (Aptenodytes forsteri) in the Ross Sea (Ainley/Ballard Antarctic field team observations at Cape Crozier for Adélie penguins for 2011/2012 season and personal communication with Gerald L. Kooyman for Emperor penguins observed in mid-1990s at Cape Washington) and also as reported by Barbosa et al. (2014) in the Antarctic Peninsula. Similar feather-loss patterns have been noted previously in African penguins (Spheniscus demersus) at rehabilitation centres and in Magellanic penguins (Spheniscus magellanicus) in colonies in Argentina (Kane et al., 2010).

Despite the wealth of information on penguin biology and ecology, there is limited information about their viruses. Most of the early studies have relied on serological assays for identifying putative paramyxoviruses, orthomyxoviruses, flavirviruses and birnaviruses in wild penguin populations (Alexander et al., 1989; Austin & Webster, 1993; Gardner et al., 1997; Miller et al., 2010; Morgan & Westbury, 1981; 1988; Morgan et al., 1981, 1985; Smith et al., 2008; Thomazelli et al., 2010) and herpesviruses and togaviruses in captive individuals (Kincaid et al., 1988; Tuttle et al., 2005). A handful of recent studies have identified some of the viruses (avipoxviruses, Newcastle disease...
viruses, adenovirus, avian influenza virus and papillomavirus) at a molecular level (Carulei et al., 2009; Hurt et al., 2014; Kane et al., 2012; Lee et al., 2014; Thomazelli et al., 2010; Varsani et al., 2014).

Our team recently identified a novel papillomavirus associated with Adélie penguins breeding at Cape Crozier, Ross Island, Antarctica (Varsani et al., 2014). Here, we report investigations at Cape Royds, a much smaller colony (~3000 vs 280 000 pairs) on the opposite coast of the island (Lyver et al., 2014), where conditions are much more severe (Dugger et al., 2014). As part of an ongoing diet study in a subcolony at Cape Royds, we had access to faecal matter from which we could identify prey hard parts as well as viral pathogens. We placed a 1 m² tray on the ground within an area where Adélie penguins nest. The tray, constructed of a hardwood frame holding a densely meshed (mesh ~0.5 mm), stainless steel hardware cloth, was placed before breeding commenced to passively collect faecal material over the 2012/2013 breeding season. Adélie penguins build their nests over the tray and defecate onto it through the course of the ~3 month breeding season. Furthermore, the penguins in their subcolonies nest at high density and thereby protect their nests from predators and competitors. Thus, faecal samples from the tray may originate from multiple penguin territories, but they are exclusively from Adélie penguins. We collected faeces from the tray at the end of the season (late January) once the nests were unoccupied.

At the end of the 2012/2013 Adélie penguin breeding season, ~50 ml of the faecal matter was transferred from the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatulas. The faecal samples were stored and transferred to the laboratory in a frozen state. Approximately half of the tray into 50 ml sterile tubes using sterile wooden spatula

Polyomaviruses are known to infect a wide range of mammalian and avian hosts. Apart from four human polyomaviruses that belong to the genus Wukipolyomavirus (WU polyomavirus, KI polyomavirus, Human polyomavirus 6 and Human polyomavirus 7), all remaining mammalian polyomaviruses belong to the genus Orthopolyomavirus. All the avian-infecting polyomaviruses belong to the genus Avipolyomavirus (Johne et al., 2011). Six avipolyomaviruses species have been identified to date, namely: Avian polyomavirus (infecting various parrot species) (Müller & Nitschke, 1986), Goose hemorrhagic polyomavirus (infecting geese and ducks) (Guerin et al., 2000), Finch polyomavirus (infecting Pyrrhula pyrrhula griseiventris), Crow polyomavirus (infecting Corvus monedula) (Johne et al., 2006), Canary polyomavirus (infecting Serinus canaria) (Halami et al., 2010) and Butcherbird polyomavirus (infecting Cracticus torquatus) (Bennett & Gillett, 2014). Avian polyomaviruses are known to cause inflammatory disease in birds; in some species the acute clinical disease can result in high mortality (Guerin et al., 2000; Johne & Müller, 2007; Krautwald et al., 1989) and in some species chronic disease of skin and feathers (Krautwald et al., 1989; Wittig et al., 2007).

Based on the 4981 nt de novo-assembled contig sequence encoding a putative T-Ag, we designed a set of abutting primers Pry-AP-F: 5’-GCATCCAAAGGCTAGGTCCAAAGCGC-3’ and Pry-AP-R: 5’-CCCAGTTGGAATTTCACGC-3’. The resulting amplicon was resolved on a 96 % agarose gel (stained with SYBR Safe DNA gel stain), excised, gel purified and cloned into the pET1.2 plasmid (ThermoFisher). The recombinant plasmid containing the amplicon was Sanger-sequenced by primer walking at Macrogen Inc. (South Korea). The Sanger-sequenced contigs were assembled using DNAbaser v.4 (Heracle BioSoft S.R.L.). All nucleotide and amino acid pairwise identities were calculated using SDT v.1.2 (Muhire et al., 2014). The resulting amplicon was resolved on a 0.7 % agarose gel (stained with SYBR Safe DNA gel stain), excised, gel purified and cloned into the pET1.2 plasmid (ThermoFisher). The recombinant plasmid containing the amplicon was Sanger-sequenced by primer walking at Macrogen Inc. (South Korea). The Sanger-sequenced contigs were assembled using DNAbaser v.4 (Heracle BioSoft S.R.L.). All nucleotide and amino acid pairwise identities were calculated using SDT v.1.2 (Muhire et al., 2014).
AdPyV (KP033140)
APyV (AF241168)
FPyV (DQ192571)
CaPyV (GU345044)
CPyV (DQ192570)
BuPyV (KF360862)
GHPyV (AY140894)

(a) (b)

Full genome
Pairwise identity (%)
100 93 87 80 73 67 60 53 47 40 33

t-Ag

T-Ag VP1 VP2 VP3
nt aa nt aa

AdPyV (KP033140)
APyV (AF241168)
FPyV (DQ192571)
CaPyV (GU345044)
CPyV (DQ192570)
BuPyV (KF360862)
GHPyV (AY140894)

(c) (d)

VP1

AdPyV (KP033140)
APyV (AF241168)
FPyV (DQ192571)
CaPyV (GU345044)
CPyV (DQ192570)
BuPyV (KF360862)
GHPyV (AY140894)

(e) (f)

VP2

AdPyV (KP033140)
APyV (AF241168)
FPyV (DQ192571)
CaPyV (GU345044)
CPyV (DQ192570)
BuPyV (KF360862)
GHPyV (AY140894)

(g) (h)

VP3

Fig. 1. (a) Genome organization of Adélie penguin polyomavirus (AdPyV). (b) Percentage pairwise identities of the full genomes of the avipolyomaviruses. (c–g) Nucleotide (nt) and amino acid (aa) percentage pairwise identities of the avipolyomavirus sequences of the large T-Antigen (T-Ag) (c), small T-Antigen (t-Ag) (d), VP1 (e), VP2 (f) and VP3 (g). APyV, Avian polyomavirus; FPyV, Finch polyomavirus; CaPyV, Canary polyomavirus; CPyV, Crow polyomavirus; BuPyV, Butcherbird polyomavirus; GHPyV, Goose hemorrhagic polyomavirus.

Identification of avian polyomavirus in Adélie penguin
Fig. 2. ML phylogenetic trees of the large T-Antigen (T-Ag) (a) and VP1 : VP2 concatenated amino acid sequences (b). The T-Ag ML phylogenetic tree was rooted with papillomavirus E1 sequences and the VP1 : VP2 tree was mid-point rooted. Branches with < 60% bootstrap support have been collapsed. See Table S1, available in the online Supplementary Material, for full names.
tentatively named this viral genome Adélie penguin polyomavirus (AdPyV). The T-Ag (1983 nt), t-Ag (537 nt), VP1 (1083 nt), VP2 (1050 nt) and VP3 (705 nt) share ~54–64% pairwise nucleotide identity and ~30–58% amino acid identity with homologous ORFs encoded by avipolyomaviruses (Fig. 1).

We also undertook posterior mapping of paired-end reads from Illumina sequencing of the 4988 nt AdPyV genome using Bowtie 2 v.2.2.3 (Langmead & Salzberg, 2012). The reads mapped across the whole genome with >250-fold coverage (Fig. S1). The de novo-assembled contig was 99.9% similar to the PCR–recovered and cloned genome that was Sanger-sequenced.

Given that there is some movement of individual penguins between colonies on Ross Island (Dugger et al., 2010), we screened the faecal sample from Cape Crozier reported by Varsani et al. (2014) and also checked the sequence reads resulting from this sample for AdPyV but found no evidence of it in the sample. Similarly, we did not find any evidence of the P. adeliae papillomavirus (Varsani et al., 2014) in the Cape Royds sample.

We assembled two datasets of amino acid sequences, the T-Ags and concatenated VP1:VP2 of representative wuki-, ortho- and avipolyomaviruses. In the T-Ag dataset, we also included the T-Ag-like sequences encoded by Bandicoot papillomatosis carcinomatosis virus 1 and 2 (BPCV-1 and -2, respectively) (Woolford et al., 2007) and Japanese eel endothelial cells-infecting virus (Mizutani et al., 2011) and selected avian papillomavirus E1 sequences (to root the tree). These datasets were aligned using MUSCLE (Edgar, 2004). The alignments were used to infer maximum-likelihood (ML) phylogenetic trees using PhyML 3.0 (Guindon et al., 2010); 1000 bootstrap replicates using the RtRev + G + I + F model of substitution (for both aligned datasets) chose the best-fit model using ProTest (Daruiba et al., 2011). The T-Ag ML phylogenetic tree was rooted with the papillomavirus E1 sequences, whereas the VP1:VP2 one was mid-point rooted. All branches with <60% bootstrap support were collapsed using Mesquite v.2.7 (http://mesquiteproject.org/). The ML phylogenetic analysis of the T-Ag clearly showed that the AdPyV clusters with other avipolyomaviruses (Fig. 2). As noted previously (Woolford et al., 2007), the T-Ag-like protein sequence encoded by BPCV-1 was phylogenetically basal to the avipolyomaviruses (Fig. 2). Within the protein sequences of the T-Ag, we identified the polyomavirus conserved region 1 (CR1; LIRLL) and hexapeptide (pRB; LYCEE), a putative nuclear localization signal (NLS; PPKSQP), a...
zinc-finger motif (CQDCKQKANTPFGKLKSKWMGGHCDDH) and ATPase motifs (GPVNSKT and GSPVPVNLE), which are all relatively conserved among all polyomaviruses (Ehlers & Moens, 2014) and BCPVs (Woolford et al., 2007) (Fig. 3). The consensus CXCX_2X_C protein phosphate 2A binding domain that is found in almost all mammalian t-Ags was not found to be encoded by AdPyV. In the VP1:VP2 concatenated amino acid sequence ML phylogenetic tree, the avipolyomavirus sequences are nested within those of the orthopolyomaviruses; VP1:VP2 of the wukipolyomaviruses are the most divergent; hence the proposal by Johne et al. (2011) to establish a new genus for these viruses. Nonetheless, the VP1 of AdPyV shared the highest pairwise amino acid identity (58%) with that of butcherbird polyomavirus and goose hemorrhagic polyomavirus (GHPyV), whereas the VP2 was most closely related to GHPyV (42%) (Fig. 1).

Based on the ML phylogenetic analysis of the T-Ag coupled with pairwise identities of the ORFs, AdPyV represents a new species of avipolyomaviruses. As highlighted by Varsani et al. (2014) for the P. adeliae papillomavirus, we are confident that AdPyV is associated with Adélie penguins and that the chances of contamination by the South Polar skua (Stercorarius maccormicki; preys on Adélie penguin eggs and chicks), the only other bird in the subcolony, are extremely slim; the skua do not nest within the faecal trays, are chased away by the penguins, are outnumbered by the penguins by >200:1, and in fact one skua pair defends ~1000 nests, and the tray, within its territory. This report expands our current knowledge of the host range of avipolyomaviruses and to the best of our knowledge this is the first report of a polyomavirus associated with penguins. The identification of this novel avipolyomavirus will enable us to design specific PCR probes to determine its prevalence and host range, given that this information is unknown; furthermore, complete genomes can be recovered from samples that test positive to determine viral diversity.

To the best of our knowledge, this is one of the five genomes of viruses associated with Antarctic animals to be characterized to date; previously identified viral genomes include those of a P. adeliae papillomavirus and an avian influenza virus subtype H1N2 from Adélie penguins (Hurt et al., 2014; Varsani et al., 2014) and adenoviruses from South Polar skuas and chinstrap penguins (Pygoscelis antarctica) (Lee et al., 2014; Park et al., 2012). This highlights the poor knowledge of viruses associated with Antarctic animals. It is worth noting that this poor knowledge of viruses in Antarctica goes beyond just Antarctic animals. It is worth noting that this poor knowledge of paramyxoviruses isolated from penguins in Antarctica and sub-Antarctica during 1976–1979. Arch Virol 109, 135–143.


Carulei, O., Douglass, N. & Williamson, A. L. (2009). Phylogenetic analysis of three genes of Penguinpox virus corresponding to Vaccinia virus G8R (VLTF-1), A3L (P4b) and H3L reveals that it is most closely related to Turkeypox virus, Ostrichpox virus and Pigeonpox virus. Virol J 6, 52.


Gerits, N. & Moens, U. (2009). A novel polyomavirus (goose hemorrhagic...
polyomavirus) is the agent of hemorrhagic nephritis enteritis of geese. *J Virol* 74, 4523–4529.


