



# Complete Genome Sequence of a Polyomavirus Recovered from a Pomona Leaf-Nosed Bat (*Hipposideros pomona*) Metagenome Data Set

Paul G. Cantalupo,<sup>a</sup> Christopher B. Buck,<sup>b</sup> James M. Pipas<sup>a</sup>

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA<sup>a</sup>; Lab of Cellular Oncology, National Cancer Institute, Bethesda, Maryland, USA<sup>b</sup>

**ABSTRACT** We report here the complete genome sequence of a polyomavirus found in a nasal/rectal metagenome of *Hipposideros pomona* (Pomona leaf-nosed bat). Interestingly, the genetic organization and phylogenetic relationships of the new virus suggest greater similarity to recently discovered fish-associated polyomaviruses rather than to polyomavirus species previously observed in bats.

Polyomaviruses are nonenveloped icosahedral viruses with a circular double-stranded DNA (dsDNA) genome about 5 kb in size and known to infect mammals and birds. Many new polyomaviruses have been discovered since the advent of next-generation sequencing technologies, and polyomavirus sequence detection has been extended to several fish and arthropod species (1–3). Polyomavirus genomes are divided into early and late-gene regions, which occupy opposing strands. The 5′ ends of the early and late regions are separated by a regulatory region (RR) containing the origin of replication. Alternatively spliced early region transcripts encode various tumor (T) antigens, including large T (LTag) and small T, while the major and minor capsid proteins (VP1 and VP2, respectively) are encoded in the late region.

We applied our custom viral discovery pipeline to the metagenomes of Chinese bats (4) and found a previously unidentified 4,398-bp contig from *Hipposideros pomona* (SRR2063883 library). A BLASTX search showed 41% identity between the contig and the VP1 protein of black sea bass polyomavirus (BassPyV1; accession no. NC\_025790). Using an in-house script, a second contig was used to construct a complete circular sequence of 4,676 bp. The complete genome of the Pomona leaf-nosed bat-associated polyomavirus 1 (PomBAPyV1) was covered at a read depth from 9 to 502×, with 9,236 total reads.

PomBAPyV1 contains a typical opposing arrangement of open reading frames (ORFs) containing LTag, VP1, and VP2 genes. Similar to fish-associated polyomaviruses, LTag appears to be encoded by a single nonspliced ORF. A BLASTP search using the predicted LTag protein sequence of PomBAPyV1 showed BassPyV1 LTag as the top hit, with 37% identity. Most polyomavirus LTags contain an HPDKGG motif that is a hallmark of DNAJ chaperone protein homologs. PomBAPyV1 LTag encodes an HPDKDP motif, and an HHpred search (5) confirms predicted structural similarity to DNAJ homologs, including the DNAJ domain of SV40 LTag (*P* value, 1.7e-7). A canonical LXCXE motif that might mediate interactions with the retinoblastoma (Rb) family of tumor suppressor proteins was found near the carboxy-terminal end of LTag. A characteristic nucleotide binding motif (GPFDCGK) was observed at LTag residues 376 to 382. No ORFs associated with small T, middle T, or alternative to large T open reading frame (ALTO) were found, again suggesting an organization similar to the fish-associated polyomaviruses (1).

Received 17 August 2016 Accepted 21 November 2016 Published 19 January 2017

**Citation** Cantalupo PG, Buck CB, Pipas JM. 2017. Complete genome sequence of a polyomavirus recovered from a Pomona leaf-nosed bat (*Hipposideros pomona*) metagenome data set. *Genome Announc* 5:e01053-16. <https://doi.org/10.1128/genomeA.01053-16>.

**Copyright** © 2017 Cantalupo 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 James M. Pipas, [pipas@pitt.edu](mailto:pipas@pitt.edu).

Phylogenetic analyses of the predicted 348-amino acid (aa) VP1 protein indicate that it occupies the same clade as VP1 proteins found in fish-associated polyomaviruses (bootstrap, 96%) (6). The predicted 511-aa VP2 protein encodes a predicted N-terminal myristoylation sequence but otherwise shows little primary sequence similarity to previously reported VP2 proteins. The only significant polyomavirus-related hit in a DELTA-BLAST search using VP2 was BassPyV1 VP2, with 22% identity and an *E* value of  $2e-10$ .

The genomic organization and phylogenetics of PomBAPyV1 indicate that it is most closely related to polyomaviruses previously observed in fish. It remains unclear whether the source of the virus is dietary/environmental or reflects productive infection of the bat itself.

**Accession number(s).** The genome sequence was deposited in ENA at [LT605004](#).

## REFERENCES

1. 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 TF, 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>.
2. Feng H, Shuda M, Chang Y, Moore PS. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319:1096–1100. <https://doi.org/10.1126/science.1152586>.
3. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R, Brennan DC, Mkakosya RS, Manary MJ, Gordon JL, Wang D. 2013. Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. *Virology* 436:295–303. <https://doi.org/10.1016/j.virol.2012.12.005>.
4. Wu Z, Ren X, Yang L, Hu Y, Yang J, He G, Zhang J, Dong J, Sun L, Du J, Liu L, Xue Y, Wang J, Yang F, Zhang S, Jin Q. 2012. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. *J Virol* 86:10999–11012. <https://doi.org/10.1128/JVI.01394-12>.
5. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.
6. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465–W469. <https://doi.org/10.1093/nar/gkn180>.