



# Draft Genome Sequence of a Novel Adenovirus Recovered from the Metagenome of Agile Wallabies

 God'spower Richard Okoh,<sup>a</sup> Ellen Ariel,<sup>a</sup> David Whitmore,<sup>a</sup>  Paul F. Horwood<sup>a</sup>

<sup>a</sup>James Cook University, College of Public Health, Medical and Veterinary Sciences, Townsville, Queensland, Australia

**ABSTRACT** Here, we report the draft genome sequence of a novel agile wallaby adenovirus that was detected in the fecal metagenome of agile wallabies. The genome is 31,512 bp long, with a G+C content of 34.4%. Currently, the pathogenic and zoonotic potential of this novel virus is unknown.

Adenoviruses are nonenveloped, icosahedral viruses with linear, unsegmented double-stranded DNA genomes (1). The genomes of adenoviruses range in size from 26 to 48 kb and commonly contain between 22 and 40 genes (1, 2). The family *Adenoviridae* is divided into six genera, namely, *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus*, *Ichtadenovirus*, and *Testadenovirus* (1, 3). Adenoviral infections do not always result in disease, although they have been associated with both single and multipathogen disease processes (4–8).

The novel Agile wallaby atadenovirus 1 (AwAdV-1) described in this report belongs to the genus *Atadenovirus* and was originally identified in the metagenome of free-ranging agile wallabies (*Notamacropus agilis*) (G. R. Okoh, E. Ariel, D. Whitmore, P. F. Horwood, submitted for publication). Briefly, five fresh fecal samples were collected from the ground at grazing sites around James Cook University and Townsville University Hospital (Townsville, Australia) in 2021. The samples were homogenized, pooled, and then virally enriched by filtration (0.25  $\mu\text{m}$ ), ultracentrifugation (100,000  $\times g$ ), and digestion with DNase I (20 U/mL). Viral DNA was then extracted using QIAamp MinElute virus kit (Qiagen). Library preparation using the Nextera DNA XT kit and Illumina sequencing (NovaSeq 6000) were performed at MacroGen (Seoul, South Korea) in paired-end 151-bp format. For this report, the sequencing reads (71,170,820) were trimmed (Trimmomatic v0.39) (9) to remove low-quality reads, normalized using BBNorm v39.01 (<https://sourceforge.net/projects/bbmap/>), and *de novo* assembled using SPAdes v3.15.5 in “careful” mode (10). The resulting contigs were searched using Diamond BLASTX (11) against the NCBI nonredundant (nr) protein database to identify the contigs corresponding to adenovirus. To assemble the genome, the reads were mapped to the identified adenoviral contig using Geneious v11.1.5 (<https://www.geneious.com>). Prediction of open reading frames (ORFs) was performed using Glimmer3 in Geneious v11.1.5, and ORF annotations were determined by conducting a BLASTX search against the NCBI nr protein databases (12). All tools were run with default parameters unless otherwise specified.

The assembled genome of AwAdV-1 was found to be 31,512 bp long, with a coverage depth of 22 $\times$  and 34.4% G+C content. The genome was predicted to contain 32 ORFs with an orientation typical of atadenoviruses. Of the 32 ORFs, 26 were annotated as having various similarities to the coding genes of other atadenoviruses (Tables 1). The IVa2, penton base protein, PX, and hexon genes showed the highest amino acid identity (71% to 87%) to the reference mammalian atadenoviruses (Table 1). The AwAdV-1 genome possesses multiple insertions and deletions in most of the genes except IVa2, pX, pVI, pVIII, and U-exon. Two fiber genes, namely, fiber and IV-1 (homologous to the fiber 2 gene in lizard adenovirus 2), were present in the genome of AWAdV-1, instead of the single long fiber gene in mammalian atadenoviruses. Phylogenetic analysis based on the full amino

**Editor** Jelle Matthijnssens, Katholieke Universiteit Leuven

**Copyright** © 2023 Okoh 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 God'spower Richard Okoh, [godspower.okoh@my.jcu.edu.au](mailto:godspower.okoh@my.jcu.edu.au), or Paul F. Horwood, [paul.horwood@jcu.edu.au](mailto:paul.horwood@jcu.edu.au).

The authors declare no conflict of interest.

**Received** 15 February 2023

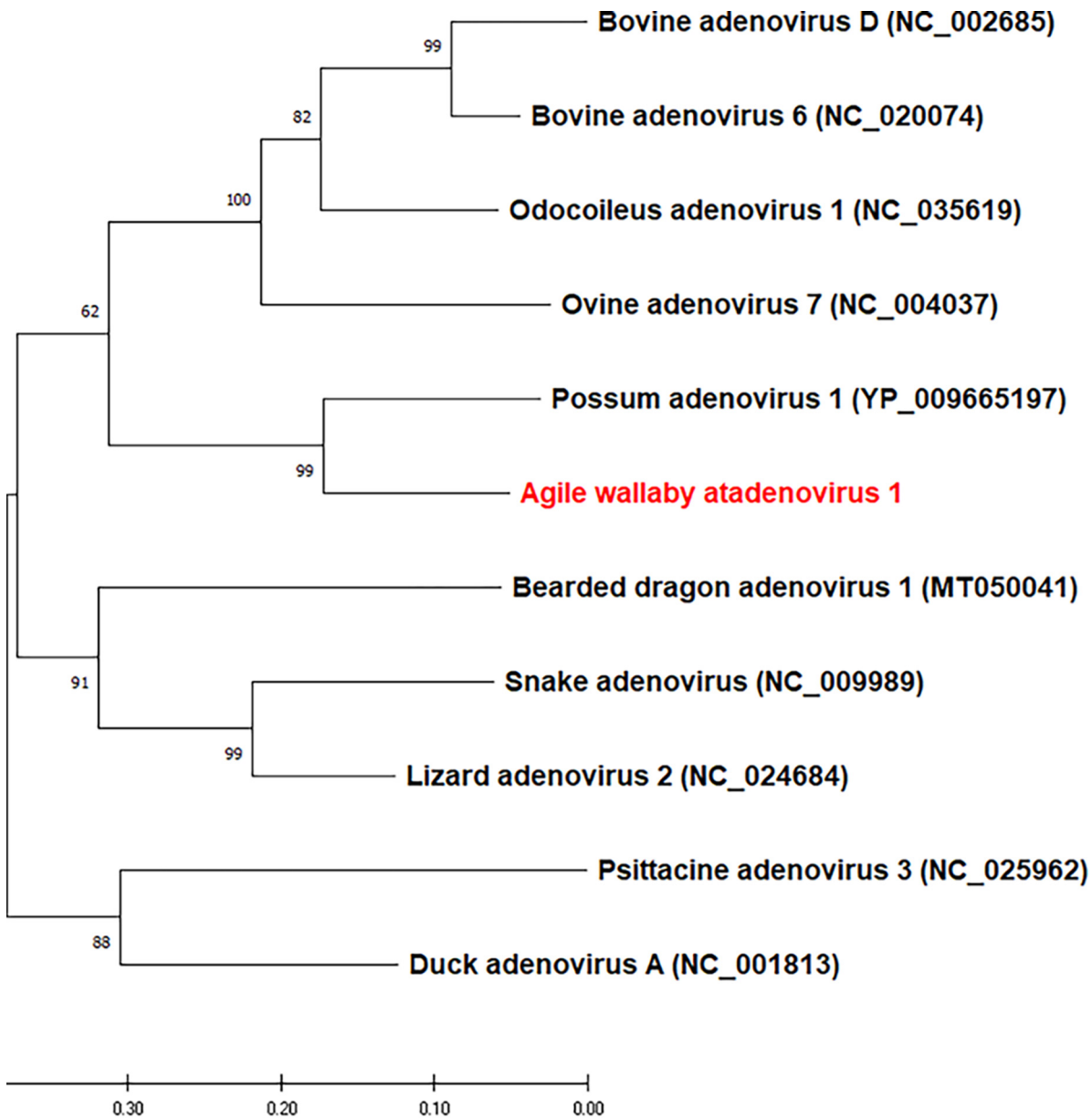
**Accepted** 26 April 2023

**Published** 18 May 2023

**TABLE 1** Sequence comparison with the genomes of mammalian adenoviruses

Data for virus (GenBank accession no.): <sup>a</sup>																				
Agile wallaby adenovirus 1		Ovine adenovirus (NC_004037)				Bovine adenovirus D (NC_002685)				Bovine adenovirus E (NC_020074)				Odocoileus adenovirus 1 (NC_035619)						
ORF	Description (putative)	Length (nt)	Length (aa)	Length (nt)	Length (aa)	Length (nt)	Length (aa)	Length (nt)	Length (aa)	Length (nt)	Length (aa)	Length (nt)	Length (aa)	Length (nt)	Length (aa)	Amino acid identity (%)	Amino acid identity (%)	Amino acid identity (%)	Amino acid identity (%)	Amino acid identity (%)
orf00001	Hypothetical protein	255	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data
orf00002	Hypothetical protein	402	134	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data
orf00003	p32K	1,032	344	861	286	819	272	45	296	891	296	40	309	930	35	40	35	35	35	35
orf00004	LH1	396	132	363	120	378	125	29	125	378	125	26	130	393	25	26	25	25	25	25
orf00005	E1B 55K	1,293	431	1,149	382	1,161	386	42	383	1,152	383	45	381	1,146	43	45	43	43	43	43
orf00006	IVa2	903	301	984	327	966	321	74	321	966	321	72	402	1,209	72	72	72	72	72	72
orf00007	Pol	3,243	1,081	3,216	1,071	3,222	1,073	58	1,073	3,222	1,073	57	1,075	3,228	58	57	58	58	58	58
orf00008	pTP	1,773	591	1,788	595	1,803	600	51	600	1,803	600	51	599	1,800	51	51	51	51	51	51
orf00009	52K	987	329	1,008	335	1,059	343	62	344	1,035	344	60	337	1,014	60	60	60	60	60	60
orf00010	pIIa	1,767	589	1,707	568	1,722	573	54	516	1,551	516	54	582	1,749	53	54	53	53	53	53
orf00011	Penton base protein	1,347	449	1,359	452	1,353	450	67	452	1,359	452	68	450	1,353	71	68	71	71	71	71
orf00012	pVII	345	115	336	111	360	119	57	118	357	118	56	117	354	51	56	51	51	51	51
orf00013	pX	102	34	216	71	216	71	81	72	219	72	0	No data	No data	No data	No data	No data	No data	No data	No data
orf00014	pVI	669	223	666	221	603	200	57	203	612	203	56	225	678	54	56	54	54	54	54
orf00015	Hexon	2,730	910	2,736	911	2,733	910	72	910	2,733	910	76	910	2,733	74	76	74	74	74	74
orf00016	23K endoprotease	606	202	606	201	606	201	60	201	606	201	61	201	606	60	61	60	60	60	60
orf00018	DNA binding protein	999	333	1,149	382	1,143	380	58	379	1,140	379	56	385	1,158	57	56	57	57	57	57
orf00020	100K	2,055	685	1,878	625	1,887	628	56	628	1,887	628	54	637	1,914	55	54	55	55	55	55
orf00021	33k	512	170	402	133	405	134	65	135	408	135	39	137	414	38	39	38	38	38	38
orf00022	pVIII	780	260	654	217	669	222	44	223	672	223	45	226	681	49	45	49	49	49	49
orf00023	U-exon	165	55	177	58	165	54	54	54	165	54	54	54	165	54	54	54	54	54	54
orf00024	Fiber	891	297	1,632	543	1,608	535	36	443	1,332	443	43	473	1,422	29	43	29	29	29	29
orf00025	IV-1	1,554	518	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data
orf00026	E4.3	600	200	714	237	654	217	38	217	657	218	38	234	705	33	38	33	33	33	33
orf00027	E4.2	678	226	663	220	678	219	42	219	660	219	42	219	678	39	42	39	39	39	39
orf00028	E4.1	441	147	429	142	429	142	37	142	429	142	38	142	429	30	38	30	30	30	30
orf00029	RH0	381	127	No data	No data	564	187	45	No data	564	187	45	No data	No data	No data	No data	No data	No data	No data	No data
orf00030	RH5	606	202	597	198	624	207	27	207	651	216	24	207	624	34	24	34	34	34	34
orf00031	Hypothetical protein	372	124	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data
orf00032	Hypothetical protein	201	67	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data
orf00033	Hypothetical protein	366	122	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data
orf00034	Hypothetical protein	372	124	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data

<sup>a</sup> aa, amino acids.



**FIG 1** Phylogenetic analysis of Agile wallaby atadenovirus 1 (shown in red), based on the amino acid sequence of penton-based protein. All the sequences used in this analysis belong to the genus *Atadenovirus*. Following multiple sequence alignment of amino acid sequences using the Muscle program in Geneious v11.1.5, a maximum likelihood phylogenetic tree was constructed using MEGA X (13) with the best model of amino acid substitution (LG + G) and 1,000 bootstrap replications.

acid sequence of penton base protein showed that AwAdV-1 belongs to the genus *Atadenovirus* and forms a distinct cluster with another marsupial adenovirus known as possum adenovirus 1 (Fig. 1).

The pathogenic potential of AwAdV-1 is unclear; however, it could be a suitable candidate for future research in vaccinology, diagnostics, and therapeutics.

**Data availability.** The raw sequence reads for this study have been deposited in the NCBI SRA database under BioProject accession number [PRJNA907146](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA907146) and BioSample accession number [SAMN31952915](https://www.ncbi.nlm.nih.gov/biosample/SAMN31952915). The novel genome sequence has been deposited at GenBank under the accession number [OQ792214](https://www.ncbi.nlm.nih.gov/genbank/OQ792214).

#### ACKNOWLEDGMENTS

We thank the Australian Wildlife Society and College of Public Health, Medical and Veterinary Sciences, James Cook University, for funding this research.

## REFERENCES

- Madarama H, Uchiyama J, Tamukai K, Katayama Y, Osawa N, Suzuki K, Mizutani T, Ochiai H. 2019. Complete genome sequence of an adenovirus-1 isolate from an African pygmy hedgehog (*Atelerix albiventris*) exhibiting respiratory symptoms in Japan. *Microbiol Resour Announc* 8:e00695-19. <https://doi.org/10.1128/MRA.00695-19>.
- Davison AJ, Benkő M, Harrach B. 2003. Genetic content and evolution of adenoviruses. *J Gen Virol* 84:2895–2908. <https://doi.org/10.1099/vir.0.19497-0>.
- Böszörményi KP, Podgorski I, Vidovszky MZ, Sós E, Benkő M, Harrach B. 2020. Full genome sequence analysis of a novel adenovirus from a captive polar bear (*Ursus maritimus*). *Virus Res* 277:197846. <https://doi.org/10.1016/j.virusres.2019.197846>.
- Jesse ST, Ciurkiewicz M, Siesenop U, Spitzbarth I, Osterhaus ADME, Baumgärtner W, Ludlow M. 2022. Molecular characterization of a bovine adenovirus type 7 (bovine atadenovirus F) strain isolated from a systemically infected calf in Germany. *Virology* 19:89. <https://doi.org/10.1186/s12985-022-01817-y>.
- Fent GM, Fulton RW, Saliki JT, Caseltine SL, Lehmkuhl HD, Confer AW, Purdy CW, Briggs RE, Loan RW, Duff GC. 2002. Bovine adenovirus serotype 7 infections in postweaning calves. *Am J Vet Res* 63:976–978. <https://doi.org/10.2460/ajvr.2002.63.976>.
- Thomson D, Meers J, Harrach B. 2002. Molecular confirmation of an adenovirus in brushtail possums (*Trichosurus vulpecula*). *Virus Res* 83:189–195. [https://doi.org/10.1016/s0168-1702\(01\)00437-3](https://doi.org/10.1016/s0168-1702(01)00437-3).
- Marschang RE. 2019. Emerging reptile viruses, p 267–273. *In* Miller RE, Lamberski N, Calle PP (ed), *Fowler's zoo and wild animal medicine: current therapy*, vol 9. Elsevier, St. Louis, MO.
- Chen EC, Yagi S, Kelly KR, Mendoza SP, Tarara RP, Canfield DR, Maninger N, Rosenthal A, Spinner A, Bales KL, Schnurr DP, Lerche NW, Chiu CY. 2011. Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. *PLoS Pathog* 7:e1002155. <https://doi.org/10.1371/journal.ppat.1002155>.
- 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>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Buchfink B, Reuter K, Drost HG. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 18:366–368. <https://doi.org/10.1038/s41592-021-01101-x>.
- Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W, Marchler-Bauer A, Landrum M, Lathrop S, Lu Z, Madden TL, O'Leary N, Phan L, Rangwala SH, Schneider VA, Skripchenko Y, Wang J, Ye J, Trawick BW, Pruitt KD, Sherry ST. 2021. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 49:D10–D17. <https://doi.org/10.1093/nar/gkaa892>.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.