



# Genome Sequences of Barmah Forest Virus Strains Isolated from Mosquitoes Trapped in Australian Defence Force Training Areas Reveal Multiple Nucleotide Insertions in the 3' Untranslated Region

Wenjun Liu,<sup>a</sup> Joanne R. Kizu,<sup>a</sup> Christina Hoare,<sup>a</sup> Ian R. Mitchell,<sup>b</sup> Penelope J. Gauci,<sup>b</sup> Aneta J. Gubala<sup>b</sup>

<sup>a</sup>Arbovirology Department, Australian Defence Force Malaria and Infectious Disease Institute, Enoggera, Queensland, Australia

<sup>b</sup>Land Division, Defence Science and Technology Group, Fishermans Bend, Victoria, Australia

**ABSTRACT** The complete genome sequences of two Barmah Forest virus (BFV) strains isolated from mosquitoes trapped in the Australian Defence Force (ADF) training areas during 2017 and 2018 reveal multiple nucleotide insertions in the 3' untranslated region (UTR) of ADF BFV strains compared with the BFV prototype strain whole-genome sequence in GenBank.

Epidemic polyarthritis (EPA) caused by Barmah Forest virus (BFV) infection is the second most frequently notified arboviral disease in Australia after Ross River virus (RRV), with approximately 1,200 cases reported annually over the last decade (1). BFV is a positive-sense, single-stranded, enveloped RNA virus in the *Alphavirus* genus of the *Togaviridae* family (2), and it is an endemic and enzootic virus in Australia that has a natural animal-mosquito-animal transmission cycle. Native animals, such as wallabies and kangaroos, are thought to be the main animals involved in the transmission cycle of infection (2). Several mosquito species, including *Culex annulirostris*, *Aedes vigilax*, *Aedes normanensis*, *Aedes notoscriptus*, and *Verrallina funerea*, were recorded to carry the virus (2, 3). Humans can be infected by spillover of virus, resulting in disease that is similar to but milder than RRV infection, having relatively benign symptoms such as fever, rash, arthralgia, and myalgia (2–5). There is no specific antiviral treatment, and no commercial BFV vaccine is available (2, 6, 7).

Full-genome sequences allow for intensive molecular evolutionary studies on vaccine development and disease control measures. Prior to our study, there was only one complete BFV genome sequence available in GenBank, that of the prototype strain BH2193, which was isolated from *C. annulirostris* mosquitoes trapped in the Barmah Forest area of Northern Victoria in 1974 (8).

We present here two BFV genome sequences obtained from the Australian Defence Force (ADF). The MIDITullyA.2017 strain was isolated from a homogenized pool of 20 *Verrallina* sp. mosquitoes captured in the ADF Tully training area (TA) (17.9°S, 145.9°E; Queensland, Australia) in 2017, while the MIDIWbTA.2018 strain was isolated from a homogenized pool of 20 *C. annulirostris* mosquitoes captured in the ADF Wide Bay TA (25.3°S, 152.8°E; Queensland, Australia) in 2018. The trapped mosquitoes were sorted according to species, placed into a 2-ml screw-cap vial with 1 ml MD (2% fetal bovine serum in RPMI 1640, 50 µg/ml penicillin/streptomycin, 50 µg/ml gentamicin, 2.5 µg/ml amphotericin B, and 10 mM HEPES) and 4 or 5 zirconium silica beads, and shaken for 1 min 30 s in a chilled block using a MiniBeadbeater-96 sample homogenizer (BioSpec Products, Bartlesville, OK, USA), followed by centrifugation (twice at 17,000 × *g* for 10 min at 4°C, with tube rotation). The mosquito homogenates were used to infect C6-36 *Aedes albopictus* cells for 3 days at 30°C as previously reported (9). Culture

**Citation** Liu W, Kizu JR, Hoare C, Mitchell IR, Gauci PJ, Gubala AJ. 2019. Genome sequences of Barmah Forest virus strains isolated from mosquitoes trapped in Australian Defence Force training areas reveal multiple nucleotide insertions in the 3' untranslated region. *Microbiol Resour Announc* 8:e00969-19. <https://doi.org/10.1128/MRA.00969-19>.

**Editor** Jelle Matthijnsens, KU Leuven

© Crown copyright 2019. 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 Wenjun Liu, [Wenjun.Liu@defence.gov.au](mailto:Wenjun.Liu@defence.gov.au).

**Received** 14 August 2019

**Accepted** 17 September 2019

**Published** 10 October 2019

**TABLE 1** Nucleotide and amino acid sequence pairwise comparisons of ADF BFV strains with the prototype BH2193 strain

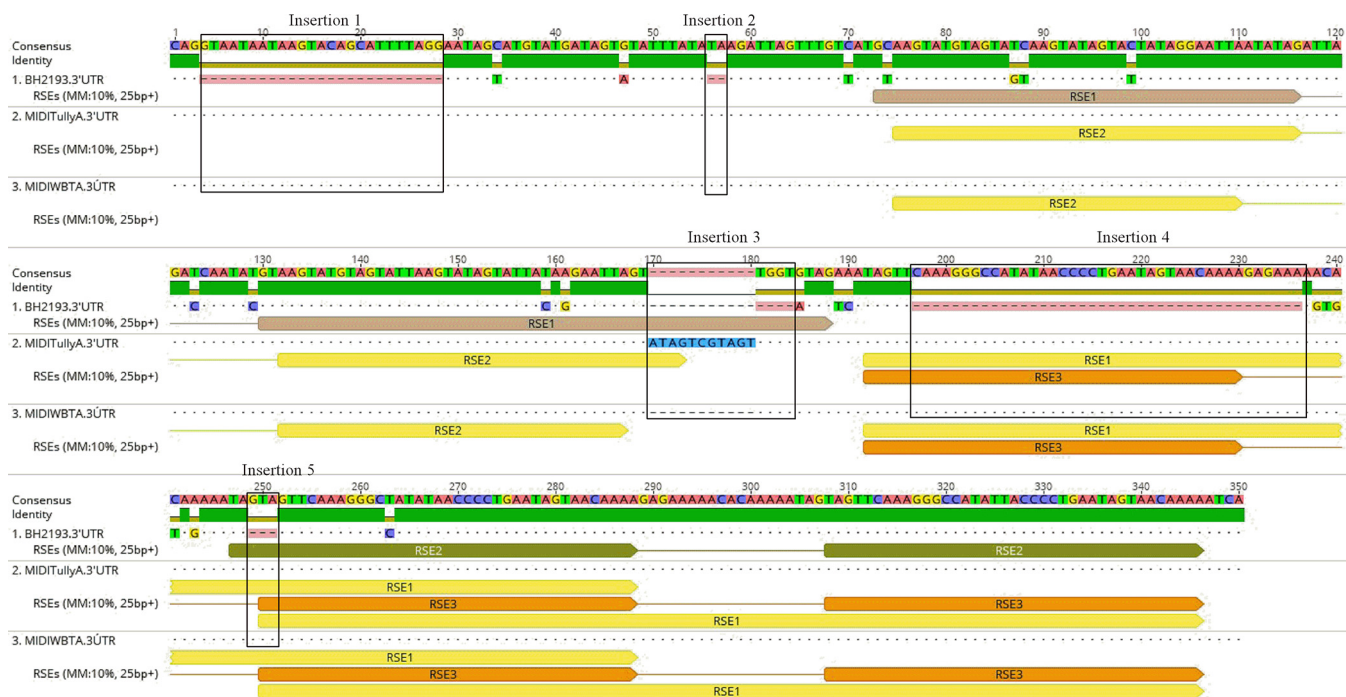
Strains compared	Similarity (%)						
	Genome	5' UTR	Nonstructural genes		Structural genes		3' UTR
			nt	aa	nt	aa	
BH2193 and MIDITullyA	97.2	95.20	98.1	99.0	98.4	99.2	80.68
BH2193 and MIDIWbTA	97.3	96.82	98.0	99.0	98.2	99.1	78.40
MIDITullyA and MIDIWbTA	99.5	98.40	99.7	99.8	99.7	99.7	97.20

supernatant was harvested, and cell debris was removed by centrifugation at  $5,000 \times g$  for 10 min. The resulting supernatant was stored at  $-80^{\circ}\text{C}$  and employed as a source of virus. Total viral RNA was extracted from a  $140\text{-}\mu\text{l}$  tissue culture supernatant using an RNeasy minikit (Qiagen) prior to being converted to cDNA using the Repli-g whole-transcriptome amplification single-cell kit (Qiagen). The cDNA library for each virus was prepared individually using the Nextera XT kit and sequenced on a MiSeq instrument using a Reagent Micro kit version 2 (300 cycles; Illumina) according to the standard protocol. The MiSeq sequence data were assembled by mapping the reads to a reference genome, that of BFV prototype strain BH2193 (GenBank accession number [NC\\_001786](#)) using Geneious software R11 (version 11.1.2). The nucleotide sequences in the 3' untranslated region (UTR) of both ADF strains were confirmed by Sanger sequencing of the reverse transcription and PCR amplicons using 5'/3' rapid amplification of cDNA ends (RACE) kit (Roche, Germany) according to the manufacturer's instructions, as described previously (9). The 3' UTR Sanger sequence results were trimmed and mapped to the BH2193 reference strain with overlapping to the E1 protein 3' nucleotide sequence end using Geneious software. The MiSeq sequence data were remapped to the generated 3' UTR sequences of both ADF strains to confirm the sequences. The open reading frames and annotations were determined by mapping each consensus sequence to the BH2193 genome using Geneious. The complete BFV sequences were aligned using ClustalW for calculating the nucleotide and the deduced amino acid differences.

In total, 1,064,168 and 774,260 paired-end reads ( $2 \times 150$  nucleotides [nt]) were sequenced for MIDITullyA.2017 and MIDIWbTA.2018, with 706,386 and 400,471 reads mapping to BFV, respectively. The resulting consensus sequence revealed that MIDITullyA.2017 and MIDIWbTA.2018 have genome lengths of 11,574 nt and 11,563 nt [without 3' poly(A)], and G+C contents of 48.5% and 48.6%, respectively. At the nucleotide level, both ADF BFV strains shared 97.2 to 97.3% similarity with the prototype BH2193 strain (Table 1). However, the similarities of the 3' UTRs are only 78.4% and 80.68%, respectively, in pairwise comparison with the prototype strain. The 99.5% nucleotide sequence similarity between MIDITullyA.2017 and MIDIWbTA.2018 isolates over a 2-year period from two different ADF training sites indicates that this virus strain is currently endemic in ADF training areas of Queensland, Australia (Table 1).

The deduced amino acid sequences of the open reading frames of both ADF BFV strains contain an additional stop codon between nonstructural nsP3 and nsP4 proteins, as reported in other alphaviruses (10, 11). In comparison with BH2193, MIDITullyA.2017 and MIDIWbTA.2018 have a total of 32 and 33 amino acid (aa) substitutions, respectively, evenly spread in nonstructural and structural proteins. It is not clear if these deviations from the prototypic sequence affect viral fitness and disease transmission and warrant further investigation.

Nucleotide alignment of the 3' UTR sequence of two ADF BFV strains with that of the prototypic strain revealed multiple insertions in the 3' UTR of both ADF strains. These insertions in both ADF strains are similar, except for insertion 3 of MIDITullyA.2017, which is 4 nt longer than that of the MIDIWbTA.2018 strain. These insertions disrupted the repeat sequence elements (RSE) (12) in the 3' UTR that were originally identified in the prototype BH2193 strain, and new RSEs are formed in both ADF strains



**FIG 1** Nucleotide alignment of 350-nt 3' UTR sequences of two ADF BFV strains with that of the prototype BH2193 strain. The repeat sequence elements (RSEs) were determined using Geneious software (version 11.1.2) with a minimum repeat length of 25 bp, a maximum mismatch of 10%, and an excluding repeat up to 10 bp longer than contained repeat. The dots indicate the consensus sequence of three BFVs, while letters in individual sequences indicate nucleotide substitutions. Dashes indicate insertions/deletions. The RSEs were annotated.

(Fig. 1). These insertions were confirmed by Sanger sequencing of reverse transcription and PCR amplicons of the 3' UTR of both ADF strains. The impact of these insertions on BFV replication and transmission is unclear. Previous studies of other alphaviral genomes have shown that deletions of RSEs in the 3' UTR region have significant effects on viral gene expression, replication, and protein translation in mosquito and avian cells, possibly through interaction with cellular proteins and microRNAs (miRNAs) (13).

**Data availability.** Raw next-generation sequencing (NGS) reads were deposited in the Sequence Read Archive under the accession numbers [SAMN11130004](https://www.ncbi.nlm.nih.gov/sra/SAMN11130004) and [SAMN11130005](https://www.ncbi.nlm.nih.gov/sra/SAMN11130005) and BioProject number [PRJNA527173](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA527173). The BFV genome sequences in this communication are publicly available in GenBank under the accession numbers [MN064696](https://www.ncbi.nlm.nih.gov/nuccore/MN064696) and [MN064697](https://www.ncbi.nlm.nih.gov/nuccore/MN064697).

## ACKNOWLEDGMENTS

This work was supported by the Joint Health Command of the Australian Defence Force.

The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The opinions expressed by the authors contributing to this article do not necessarily reflect the opinions of the institutions with which the authors are affiliated.

## REFERENCES

1. Australian Government, Department of Health, National Notifiable Diseases Surveillance System. 2019. Number of notifications of Barmah Forest virus infection, Australia, in the period of 1991 to 2018 and year-to-date notifications for 2019. Australian Government, Department of Health, Canberra, Australia. [http://www9.health.gov.au/cda/source/rpt\\_3.cfm](http://www9.health.gov.au/cda/source/rpt_3.cfm).
2. Jacups SP, Whelan PI, Currie BJ. 2008. Ross River virus and Barmah Forest virus infections: a review of history, ecology, and predictive models, with implications for tropical northern Australia. *Vector Borne Zoonotic Dis* 8:283–297. <https://doi.org/10.1089/vbz.2007.0152>.
3. Jeffery JA, Kay BH, Ryan PA. 2006. Role of *Verrallina funerea* (Diptera: Culicidae) in transmission of Barmah Forest virus and Ross River virus in coastal areas of eastern Australia. *J Med Entomol* 43:1239–1247. <https://doi.org/10.1093/jmedent/43.6.1239>.
4. Flexman JP, Smith DW, Mackenzie JS, Fraser JR, Bass SP, Hueston L, Lindsay MD, Cunningham AL. 1998. A comparison of the diseases caused

- by Ross River virus and Barmah Forest virus. *Med J Aust* 169:159–163. <https://doi.org/10.5694/j.1326-5377.1998.tb116019.x>.
5. Cashman P, Hueston L, Durrheim D, Massey P, Doggett S, Russell RC. 2008. Barmah Forest virus serology: implications for diagnosis and public health action. *Commun Dis Intell Q Rep* 32:263–266.
  6. NSW Health. 2002. Factsheet: Barmah Forest virus infection. *N S W Public Health Bull* 13:170. <https://doi.org/10.1071/NB02069>.
  7. Carey BD, Bakovic A, Callahan V, Narayanan A, Kehn-Hall K. 2019. New World alphavirus protein interactomes from a therapeutic perspective. *Antiviral Res* 163:125–139. <https://doi.org/10.1016/j.antiviral.2019.01.015>.
  8. Dalgarno L, Short NJ, Hardy CM, Bell JR, Strauss JH, Marshall ID. 1984. Characterization of Barmah Forest virus: an alphavirus with some unusual properties. *Virology* 133:416–426. [https://doi.org/10.1016/0042-6822\(84\)90407-0](https://doi.org/10.1016/0042-6822(84)90407-0).
  9. Liu W, Pickering P, Duchene S, Holmes EC, Aaskov JG. 2016. Highly divergent dengue virus type 2 in traveler returning from Borneo to Australia. *Emerg Infect Dis* 22:2146–2148. <https://doi.org/10.3201/eid2212.160813>.
  10. Li GP, Rice CM. 1989. Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational read-through and its effect on virus replication. *J Virol* 63:1326–1337.
  11. Myles KM, Kelly CL, Ledermann JP, Powers AM. 2006. Effects of an opal termination codon preceding the nsP4 gene sequence in the O’Nyong-Nyong virus genome on *Anopheles gambiae* infectivity. *J Virol* 80:4992–4997. <https://doi.org/10.1128/JVI.80.10.4992-4997.2006>.
  12. Lee E, Stocks C, Lobigs P, Hislop A, Straub J, Marshall I, Weir R, Dalgarno L. 1997. Nucleotide sequence of the Barmah Forest virus genome. *Virology* 227:509–514. <https://doi.org/10.1006/viro.1996.8343>.
  13. Hyde JL, Chen R, Trobaugh DW, Diamond MS, Weaver SC, Klimstra WB, Wilusz J. 2015. The 5’ and 3’ ends of alphavirus RNAs—non-coding is not non-functional. *Virus Res* 206:99–107. <https://doi.org/10.1016/j.virusres.2015.01.016>.