

1 **Genome sequences of equid herpesviruses 2 and 5**

2

3 Gavin S. Wilkie,^a Karen Kerr,^a James P. Stewart,^b Michael J. Studdert,^c Andrew J. Davison^a#

4

5 MRC – University of Glasgow Centre for Virus Research, Glasgow, United Kingdom^a;

6 Department of Infection Biology, Institute of Infection and Global Health, University of

7 Liverpool, Liverpool, United Kingdom^b; Centre for Equine Virology, School of Veterinary Science,

8 The University of Melbourne, Australia^c

9

10 Running title: EHV2 and EHV5 sequences

11

12 #Address correspondence to Andrew J. Davison, andrew.davison@glasgow.ac.uk

13 **ABSTRACT**

14 We have resequenced the genome of equid herpesvirus 2 (EHV2) strain 86/67, and sequenced
15 the genomes of EHV2 strain G9/92 and equid herpesvirus 5 (EHV5) strain 2-141/67. The most
16 prominent genetic differences are the dissimilar locations of the IL-10-like genes and the
17 presence of an OX-2-like gene only in EHV5.

18

19 **GENOME ANNOUNCEMENT**

20 Equid herpesviruses 2 and 5 (EHV2 and EHV5) belong to genus *Percavirus*, subfamily
21 *Gammaherpesvirinae*, family *Herpesviridae* (1, 2). Evidence for causation in the diseases
22 associated with these viruses remains elusive (3), except in the case of EHV5-induced
23 pulmonary fibrosis (4). The genome sequence of EHV2 86/67 (5, 6) has been published (7). We
24 have resequenced this strain, and have also sequenced EHV2 G9/92 (8) and EHV5 2-141/67 (6).

25 Paired-end, 250 nucleotide reads were generated from the EHV2 86/67 and EHV5 2-
26 141/67 DNA preparations analyzed originally (1), and from EHV2 G9/92 DNA also isolated at
27 that time, by using an Illumina MiSeq (v2 chemistry). The EHV2 86/67 sequence in GenBank
28 accession no. U20824.1 (7) was corrected by making directed assemblies using BWA (9) and
29 Tanoti (<http://www.bioinformatics.cvr.ac.uk/Tanoti/index.php>), and viewing them by using
30 Tablet (10). Draft EHV2 G9/92 and EHV5 2-141/67 sequences were assembled *de novo* by using
31 ABySS (11), improved by using GapFiller (12) and iCORN2 (13), and assessed by making directed
32 assemblies. Regions of low coverage or containing repeats were assessed by PCR amplification
33 and Sanger sequencing. The EHV2 G9/92 and EHV5 2-141/67 genome termini were identified
34 by using a PCR-based method (14).

35 For EHV2 86/67, EHV2 G9/92, and EHV5 2-141/67, respectively, 1,617,084, 7,239,808, and
36 2,411,370 reads were obtained, the majority (84, 91, and 94%) aligning (using Tanoti) with the
37 sequences at coverage values of 1812, 8776, and 3085 reads per nucleotide. The genome sizes
38 are 184,439, 186,110, and 182,380 bp, respectively, including an unmatched, complementary
39 nucleotide at the 3'-end of each DNA strand. The slightly larger size of the EHV5 2-141/92

40 genome from that estimated by restriction site analysis (179 kbp) (15) is due to the absence of a
41 6176 bp region from the maps. The EHV2 86/67 and EHV2 G9/92 genomes contain a terminal
42 direct repeat (TR) of 17,553 and 18,332 bp, respectively. TR is much smaller (10 bp) in the EHV5
43 2-141/67 genome, in support of previous evidence (15).

44 The number of functional open reading frames (ORFs) was estimated conservatively to be
45 78 in EHV2 (one duplicated in TR) and 79 in EHV5. Six ORFs (E1, E5A, E6A, ORF51, ORF74, and
46 E9) exhibit <80% amino acid sequence identity between the EHV2 strains. Diversity in EHV2 has
47 been studied previously by restriction site analysis (16), and that of two ORFs in the list above
48 (E1 and ORF74) by sequencing (17). Seven EHV5 ORFs (E1, E3, E6A, ORF27, ORF51, E9, and E10)
49 exhibit <40% amino acid sequence identity to their EHV2 orthologs. One EHV2 ORF (E7,
50 encoding an IL-10-like protein) occupies a dissimilar genome location in EHV5. One EHV2 ORF
51 (E6C) lacks an EHV5 ortholog, and two EHV5 ORFs (E6B and E11, the latter encoding an OX-2-
52 like protein) lack EHV2 orthologs. The largest divergent regions in the EHV5 genome are the 15
53 kbp at the left terminus and the 20 kbp at the right terminus.

54

55 **Nucleotide sequence accession numbers.** The EHV2 G9/92 and EHV5 2-141/67 genome
56 sequences have been deposited in GenBank under the accession nos. KM924294 and
57 KM924295, respectively.

58

59 **ACKNOWLEDGMENTS**

60 This work was supported by the UK Medical Research Council (grant number MC_UU_12014/3).

61 We thank Neil Edington (then at the Royal Veterinary College, London) for providing EHV2
62 G9/92, and Wai Kwong Lee and Andrew Carswell (BHF Glasgow Cardiovascular Research Centre,
63 University of Glasgow, Glasgow) for providing Sanger sequencing services.

64

65 **REFERENCES**

- 66 1. **Telford EAR, Studdert MJ, Agius CT, Watson MS, Aird HC, Davison AJ.** 1993. Equine
67 herpesviruses 2 and 5 are γ -herpesviruses. *Virology* **195**:492-499.
- 68 2. **Pellett PE, Davison AJ, Eberle R, Ehlers B, Hayward GS, Lacoste V, Minson AC, Nicholas J,**
69 **Roizman B, Studdert MJ, Wang F.** 2011. *Herpesviridae*, p 111-122. In King AMQ, Adams MJ,
70 Carstens EB, Lefkowitz EJ (ed), *Virus taxonomy*, ninth report of the International Committee
71 on Taxonomy of Viruses. Elsevier Academic Press, London, United Kingdom.
- 72 3. **Hartley CA, Dynon KJ, Mekuria ZH, El-Hage CM, Holloway SA, Gilkerson JR.** 2013. Equine
73 gammaherpesviruses: perfect parasites? *Vet Microbiol* **167**:86-92.
- 74 4. **Williams KJ, Robinson NE, Lim A, Brandenberger C, Maes R, Behan A, Bolin SR.** 2013.
75 Experimental induction of pulmonary fibrosis in horses with the gammaherpesvirus equine
76 herpesvirus 5. *PLoS One* **10**:e77754.
- 77 5. **Studdert MJ, Turner AJ, Peterson JE.** 1970. Equine herpesviruses. 1. Isolation and
78 characterisation of equine rhinopneumonitis virus and other equine herpesviruses from
79 horses. *Aust Vet J* **46**:83-89.
- 80 6. **Browning GF, Studdert MJ.** 1987. Genomic heterogeneity of equine betaherpesviruses. *J*
81 *Gen Virol* **68**:1441-1447.
- 82 7. **Telford EAR, Watson MS, Aird HC, Perry J, Davison AJ.** 1995. The DNA sequence of equine
83 herpesvirus 2. *J Mol Biol* **249**:520-528.
- 84 8. **Welch HM, Bridges CG, Lyon AM, Griffiths L, Edington N.** 1992. Latent equid herpesviruses
85 1 and 4: detection and distinction using the polymerase chain reaction and co-cultivation
86 from lymphoid tissues. *J Gen Virol* **73**:261-268.
- 87 9. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler
88 transform. *Bioinformatics* **26**:589-595.
- 89 10. **Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D.** 2013.
90 Using Tablet for visual exploration of second-generation sequencing data. *Briefings in*
91 *Bioinformatics* **14**:193-202.

- 92 11. **Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I.** 2009. ABySS: a parallel
93 assembler for short read sequence data. *Genome Res* **19**:1117-1123.
- 94 12. **Boetzer M, Pirovano W.** 2012. Toward almost closed genomes with GapFiller. *Genome Biol*
95 **13**:R56.
- 96 13. **Otto TD, Sanders M, Berriman M, Newbold C.** 2010. Iterative Correction of Reference
97 Nucleotides (iCORN) using second generation sequencing technology. *Bioinformatics*
98 **26**:1704-1707.
- 99 14. **Wilkie GS, Davison AJ, Watson M, Kerr K, Sanderson S, Bouts T, Steinbach F, Dastjerdi A.**
100 2013. Complete genome sequences of elephant endotheliotropic herpesviruses 1A and 1B
101 determined directly from fatal cases. *J Virol* **87**:6700-6712.
- 102 15. **Agius CT, Nagesha HS, Studdert MJ.** 1992. Equine herpesvirus 5: comparisons with EHV2
103 (equine cytomegalovirus), cloning, and mapping of a new equine herpesvirus with a novel
104 genome structure. *Virology* **191**:176-186.
- 105 16. **Browning GF, Studdert MJ.** 1989. Physical mapping of the genomic heterogeneity of
106 isolates of equine herpesvirus 2 (equine cytomegalovirus). *Arch Virol* **104**:87-94.
- 107 17. **Sharp EL, Farrell HE, Borchers K, Holmes EC, Davis-Poynter NJ.** 2007. Sequence analysis of
108 the equid herpesvirus 2 chemokine receptor homologues E1, ORF74 and E6 demonstrates
109 high sequence divergence between field isolates. *J Gen Virol* **88**:2450-2462.