1 **Genome sequence of white sturgeon herpesvirus 1 isolated from farmed white sturgeon**

2 **(*Acipenser transmontanus*)**

3

4 Thomas B. Waltzek[a,b,#], Kuttichantran Subramaniam[a], Andor Doszpoly[c], Joseph Hughes[d], Matej

5 Vučak[d], Andrew J. Davison[d]

6

7 [a]Department of Infectious Diseases and Immunology, College of Veterinary Medicine, University of

8 Florida, Gainesville, Florida, USA

9 [b]Present address: Washington Animal Disease Diagnostic Laboratory and Department of

10 Veterinary Microbiology and Pathology, College of Veterinary Medicine, Washington State

11 University, Pullman, Washington, USA

12 [c]Veterinary Medical Research Institute, Budapest, Hungary

13 [d]MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom

14

15 Running title: Genome sequence of white sturgeon herpesvirus 1

16

17 #Address correspondence to Thomas Waltzek, tomwaltzek@gmail.com

ABSTRACT

The genome sequence of white sturgeon herpesvirus 1, which was isolated from farmed white sturgeon (*Acipenser transmontanus*), was determined. Comparative analyses suggest the classification of this virus as a new species in a new genus in the family *Alloherpesviridae*.

Words = 39

Keywords: *Alloherpesviridae*, *Acipenser transmontanus*, sturgeon, genome.

ANNOUNCEMENT

White sturgeon herpesvirus 1 (WSHV1) was isolated on the white sturgeon skin (WSSK-1) cell line from juvenile white sturgeon experiencing elevated mortality on a commercial farm in California, USA in 1991 ([1]). Additional herpesviruses were isolated subsequently from farmed white sturgeon experiencing elevated morbidity and mortality ([2]). Based on differences from WSHV1 in serological properties and growth characteristics in cell culture, these isolates were designated white sturgeon herpesvirus 2 (WSHV2). Phylogenetic analyses based on 12 core genes present in all members of the family *Alloherpesviridae* revealed that WSHV1 is distantly related to WSHV2 ([3]) and that its closest relative is lake sturgeon herpesvirus (LSHV) ([4]).

The WSHV1 isolate (strain UC Davis) was passaged six times in WSSK-1 cells ([1]), and DNA was purified from the culture medium using a Blood & Cell Culture DNA mini kit (Qiagen). A sequencing library prepared using a Nextera XT kit (Illumina) was sequenced on a MiSeq instrument (Illumina) using a v3 600-cycle reagent kit, yielding 11,301,304 300 nt paired-end reads. The data were analyzed using default options except where stated otherwise. The reads were trimmed using Trim Galore v.0.6.6 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with the -q 25, --length 50 and --paired options. The trimmed reads were deduplicated using FastUniq v.1.1 ([5]) and subsampled using Seqtk v.1.3 (https://github.com/lh3/seqtk) with the sample, -s34 and 0.03 (to discard 97% of reads) options. The subsampled reads were assembled into contigs using SPAdes v3.14.1 ([6]) with the --careful, --cov-cutoff auto and -k 33,77,127 options. Contigs related to the genomes of members of the family *Alloherpesviridae* were identified by using BLASTX ([7]) against the NCBI non-redundant protein database. These contigs were joined to others by extracting reads

48 that extended their ends until overlaps were found. The genome terminI were assigned as

49 outlined below. The sizes of seven tandem reiterations were not resolved. The integrity of the

50 sequence was verified by aligning the trimmed reads using Bowtie 2 v2.3.1 (8) and Samtools v1.3

51 (9) and inspecting the alignment using Tablet v1.21.02.08 (10). The alignment incorporated 98% of

52 reads at an average coverage depth of 14,704 reads/nt.

53   The WSHV1 genome (195,072 bp; 40% G+C) consists of a long unique region (U_L; 145,962 bp)

54 flanked by an inverted repeat (TR_L/IR_L; 26 bp), linked to a short unique region (U_S; 42,630 bp) also

55 flanked by an inverted repeat (TR_S/IR_S; 3,214 bp), yielding the overall structure

56 $TR_L–U_L–IR_L–IR_S–U_S–TR_S$ (Fig. 1). The left terminus was assigned from a large set of reads sharing an

57 end, and the right terminus from the adjacent sequence at the IR_L-IR_S junction. A total of 129 open

58 reading frames (ORFs) encoding functional proteins were predicted using approaches described

59 previously (4) (Fig. 1).

60   Although WSHV1 and LSHV (4) are each other's closest relative, the core genes are only 76.8%

61 identical. This suggests the creation of a new genus in the family *Alloherpesviridae* incorporating

62 two new species. A second partially sequenced lake sturgeon herpesvirus (11) is more closely

63 related to LSHV than is WSHV1; the core genes are 84.3% identical.

64

65 Words = 497

66

67 **Data availability**

68   The WSHV1 genome sequence is available in GenBank under accession number OR001786. The

69 sequence reads are available under BioProject accession number PRJNA1004073.

70

76

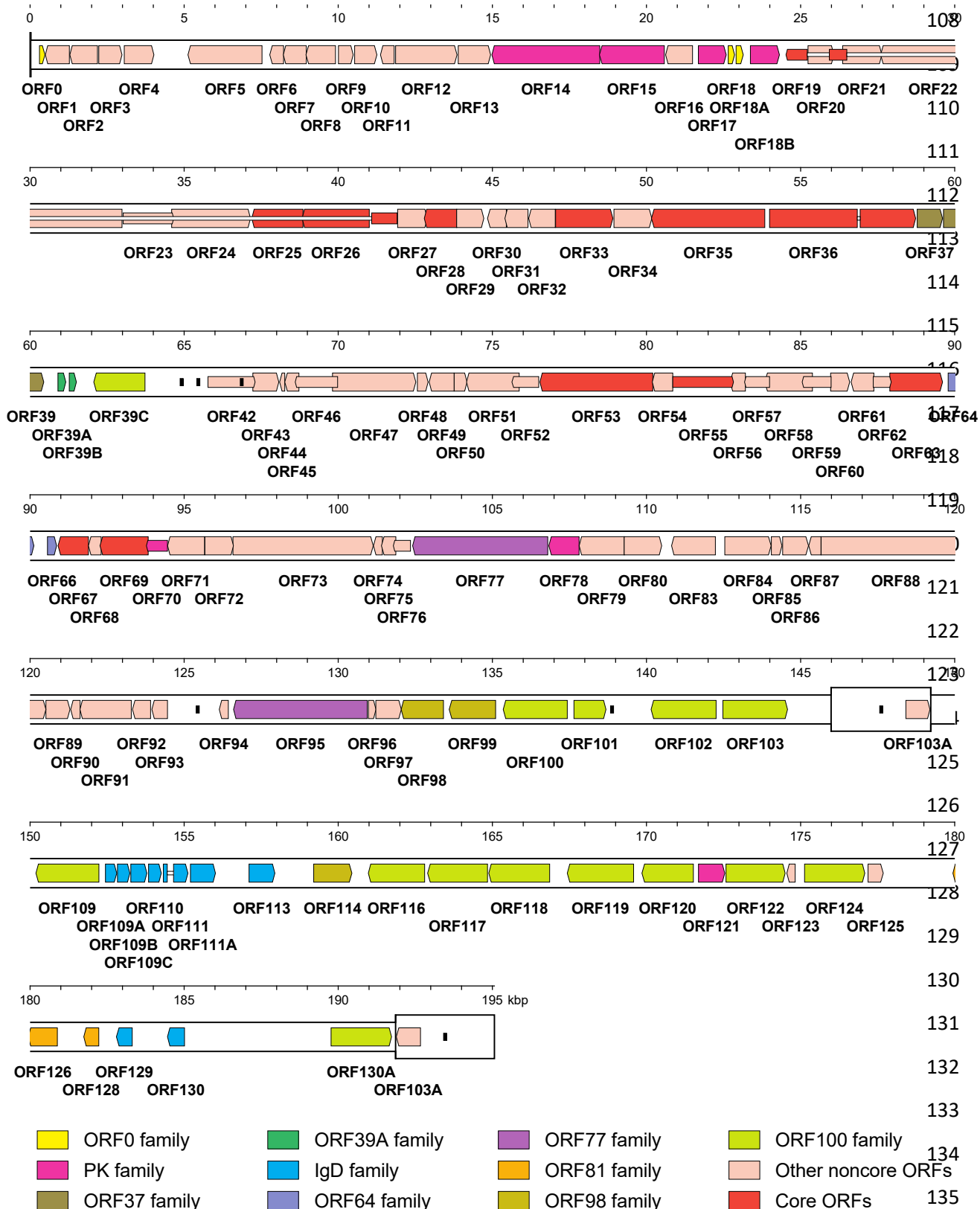77 **REFERENCES**

78   1. Hedrick RP, McDowell TS, Groff JM, Yun S, Wingfield WH. 1991. Isolation of an epitheliotropic
79   herpesvirus from white sturgeon *Acipenser transmontanus*. *Dis Aquat Org* 11:49–56.

80   2. Watson LR, Yun SC, Groff JM, Hedrick RP. 1995. Characteristics and pathogenicity of a novel
81   herpesvirus isolated from adult and subadult white sturgeon *Acipenser transmontanus*. *Dis*
82   *Aquat Org* 22:199–210.

83   3. Waltzek TB, Kelley GO, Alfaro ME, Kurobe T, Davison AJ, Hedrick RP. 2009. Phylogenetic
84   relationships in the family *Alloherpesviridae*. *Dis Aquat Org* 84:179–194.

85   4. Walker L, Subramaniam K, Viadanna PHO, Vann JA, Marcquenski S, Godard D, Kieran E, Frasca
86   S Jr, Popov VL, Kerr K, Davison AJ, Waltzek TB. 2022. Characterization of an alloherpesvirus
87   from wild lake sturgeon *Acipenser fulvescens* in Wisconsin (USA). *Dis Aquat Org* 149:83–96.

88   5. Xu H, Luo X, Qian J, Pang X, Song J, Qian G, Chen J, Chen S. 2012. FastUniq: a fast *de novo*
89   duplicates removal tool for paired short reads. *PLoS One* 7:e52249.

90   6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI,
91   Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.
92   2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.
93   *J Comput Biol* 19:455–477.

94   7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J*
95   *Mol Biol* 215:403–410.

96   8. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
97   9:357–359.

98   9. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000
99   Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and
100   SAMtools. *Bioinformatics* 25:2078–20792.

101   10. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D. 2013. Using
102   Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform* 14:193–
103   202.

104   11. Johnston AE, Shavalier MA, Scribner KT, Soto E, Griffin MJ, Waldbieser GC, Richardson BM,
105   Winters AD, Yun S, Baker EA, Larson DL, Kiupel M, Loch TP. 2022. First isolation of a
106   herpesvirus (family *Alloherpesviridae*) from Great Lakes lake sturgeon (*Acipenser fulvescens*).
107   *Animals* 12:3230.