

Genome sequence of seven human herpes simplex virus 2 (HSV2) clinical isolates from Finland

Christopher D. Bowen,^{1,2} Daniel W. Renner,^{1,2} Henrik Paavilainen,³ Tytti Vuorinen,^{3,4} Veijo Hukkanen,³ Moriah L. Szpara^{1,2,5}

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We sequenced seven herpes simplex virus 2 (HSV2) clinical isolates circulating between 2000 and 2014 in Finland. Few HSV2 virus isolates have been sequenced from Europe, particularly from Nordic countries. These sequences expand our knowledge of the global sequence diversity of HSV2 sequences and help elucidate total variation between geographic regions.

KEYWORDS herpes simplex virus 2, genomes, human herpesviruses, clinical microbiology

Herpes simplex virus 2 (HSV2; *Simplexvirus humanalpha2*, subfamily *Alphaherpesvirinae*, family *Herpesviridae*) is a double-stranded DNA virus and the causative agent of recurrent genital lesions. Seven isolates were obtained from anonymous clinical diagnostic samples of herpes lesions collected between 2000 and 2014 (Table 1) (1). Approval for the study of anonymous HSV isolates was provided by the Hospital District of South West Finland, currently designated as the Wellbeing Services County of Southwest Finland (permit number J10/17). An immunoperoxidase rapid culture assay (2) was used to type the viruses as HSV2, with confirmation by a type-specific HSV-DNA-PCR (3). Viruses were initially propagated on Vero cells (ATCC). For viral genomic DNA isolation, viral stock collected from the first or second passage on Vero cells was used to infect $\sim 1 \times 10^8$ HaCaT cells (Department of Dentistry, University of Turku [4]), and the infection was allowed to proceed to completion at 35°C (1–3 days).

Viral genomic DNA was isolated as previously described (5, 6), with Freon extraction to disrupt cellular membranes, density gradient separation of capsids, and lysis with detergent and proteinase K to free viral nucleocapsid DNA. Viral DNA was extracted with phenol-chloroform, ethanol-precipitated, and quantitated by Qubit (Thermo Fisher). Input DNA (200 ng) was quantified by Qubit, sheared using a Covaris M220, and prepared for sequencing using the TruSeq Nano DNA Low Throughput Sample Preparation Guide (Illumina). Sequencing libraries were quality-controlled using Qubit, Bioanalyzer (Agilent Technologies), and qPCR (KAPA Biosystems). Libraries were loaded on an Illumina MiSeq for 2× 300 bp paired-end sequencing.

Viral genomes were assembled using our published bioinformatics workflow (installation and instructions at <http://virga.readthedocs.org/> [7]). The VirGA-specific scripts `maf_net.py` and `compare_genomes.py` are archived at <https://bitbucket.org/szparalab/virga/src/master/> (7). Quality control utilized FastX-Toolkit and Trimmomatic (default parameters) for adapter trimming and removal of sequencing artifacts and low-quality bases (8, 9). We used Bowtie 2 to remove any DNA sequences likely derived from HaCaT host cell DNA (using [GCF_000001405.13](https://www.ncbi.nlm.nih.gov/nuccore/000001405.13) as a proxy) (7, 10). *De novo* assembly utilized SSAKE v3.8 (11), Celera 8.1 (12), and GapFiller 1.10 (insert_deviation 0.2) (13), followed by Mugsy 2.3 (14) and the VirGA script `maf_net.py` (7) to create a consensus genome for each strain from the *de novo* assembled contigs. Homology-based transfer of annotation from HSV2 strain HG52 ([NC_001798](https://www.ncbi.nlm.nih.gov/nuccore/001798) [15]) to each new consensus

Editor Jelle Matthijnsens, Katholieke Universiteit Leuven, Leuven, Belgium

Address correspondence to Moriah L. Szpara, moriah@psu.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 21 October 2025

Accepted 10 April 2026

Published 12 May 2026

Copyright © 2026 Bowen 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/).

TABLE 1 Sequence data for seven HSV2 genomes from Finland

HSV2 isolate name	Genome assembly length (bp)	# of HSV-specific paired reads	Average coverage depth	GenBank accession ID	SRA accession ID
HSV2-H1227	152,782	3,764,008	1,599×	KY922721	SRX30815609
HSV2-H1229	154,136	3,119,962	1,195×	KY922722	SRX30815610
HSV2-H12211	155,111	3,605,960	1,168×	KY922725	SRX30815611
HSV2-H12212	155,788	2,866,398	1,039×	KY922726	SRX30815612
HSV2-H1226	152,981	2,864,842	1,032×	KY922720	SRX30815608
HSV2-H1526	154,678	1,375,572	532×	KY922724	SRX30815614
HSV2-H1421	148,702	196,780	35×	KY922723	SRX30815613

genome utilized the VirGA script `compare_genomes.py` (7). The length, coverage depth, and GenBank accession IDs for these genomes are listed in Table 1. HSV2 isolate H1421 had lower sequence coverage as a result of higher host DNA contamination, leading to a partial genome assembly with incomplete termini (Table 1).

The G + C content for these viral genomes ranged from 69.1% to 70.3%. Their overall relatedness to classic HSV2 strains is shown in Fig. 1. These genome sequences have already been proven useful in reconstructing the evolutionary history of HSV2 vs the

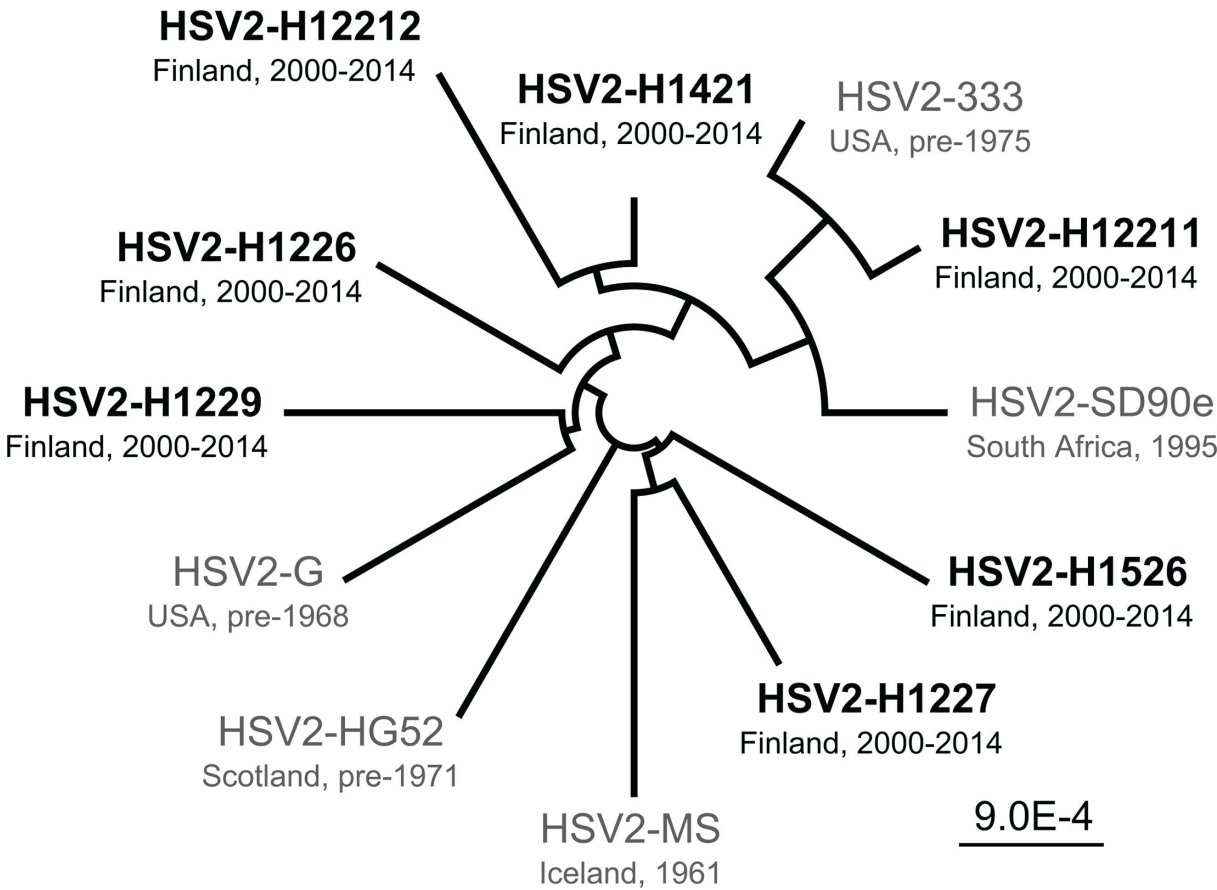


FIG 1 Genetic relatedness of Finnish HSV2 genomes. A neighbor-joining tree built based on a whole-genome alignment of seven Finnish HSV2 strains (bold) vs previously characterized HSV2 comparison strains HG52 ([NC_001798](#) [15]), MS ([MK855052](#) [18]), G ([OM370995](#) [17]), SD90e ([KF781518](#) [21]), and 333 ([LS480640](#) [22]). We used MAFFT (7.490, FFT-NS-2 with default parameters [23]) to create a whole-genome alignment. A tree generated from this full-genome alignment using the neighbor-joining method (Jukes-Cantor model, Geneious Prime 2022.2.1) illustrated their overall relatedness. These seven Finnish and five comparison HSV2 genomes shared 95.1% pairwise nucleotide sequence identity (Geneious Prime 2022.2.1).

distantly related species HSV1 (16) for comparing sequence divergence in HSV2 (17, 18) and for research on new antiviral therapeutics and diagnostic approaches (19, 20).

ACKNOWLEDGMENTS

Funding from the Academy of Finland (grant #259725) is acknowledged (V.H.). This project was supported in part by startup funds from Pennsylvania State University, a grant from the Pennsylvania Department of Health (PA DoH) CURE program, and NIH R01AI132692 and R01AI163217 (M.L.S.). The PA DoH specifically disclaims responsibility for any analyses, interpretations, or conclusions.

AUTHOR AFFILIATIONS

¹Department of Biology, Pennsylvania State University, University Park, Pennsylvania, USA

²Center for Infectious Disease Dynamics, Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, Pennsylvania, USA

³Institute of Biomedicine, University of Turku, Turku, Finland

⁴Department of Clinical Virology, Turku University Hospital, Turku, Finland

⁵Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania, USA

AUTHOR ORCIDS

Moriah L. Szpara  <http://orcid.org/0000-0001-9859-1678>

FUNDING

Funder	Grant(s)	Author(s)
Academy of Finland		Veijo Hukkanen
Pennsylvania State University		Moriah L. Szpara
Pennsylvania Department of Health	CURE	Moriah L. Szpara
National Institutes of Health	R01AI132692, R01AI163217	Moriah L. Szpara

AUTHOR CONTRIBUTIONS

Christopher D. Bowen, Conceptualization, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review and editing | Daniel W. Renner, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – review and editing | Henrik Paavilainen, Conceptualization, Formal analysis, Investigation, Resources, Writing – review and editing | Tytti Vuorinen, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing | Veijo Hukkanen, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing | Moriah L. Szpara, Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

All HSV2 clinical isolate genomes have been deposited at GenBank ([KY922720-KY922726](#)), and the raw reads have been submitted to the SRA ([SRX30815609](#), [SRX30815610](#), [SRX30815611](#), [SRX30815612](#), [SRX30815608](#), [SRX30815614](#), [SRX30815613](#)).

REFERENCES

- Bowen CD, Paavilainen H, Renner DW, Palomäki J, Lehtinen J, Vuorinen T, Norberg P, Hukkanen V, Szpara ML. 2019. Comparison of herpes simplex virus 1 strains circulating in Finland demonstrates the uncoupling of whole-genome relatedness and phenotypic outcomes of viral infection. *J Virol* 93:e01824-18. <https://doi.org/10.1128/JVI.01824-18>
- Ziegler T, Waris M, Rautiainen M, Arstila P. 1988. Herpes simplex virus detection by macroscopic reading after overnight incubation and immunoperoxidase staining. *J Clin Microbiol* 26:2013–2017. <https://doi.org/10.1128/jcm.26.10.2013-2017.1988>
- Hukkanen V, Rehn T, Kajander R, Sjöroos M, Waris M. 2000. Time-resolved fluorometry PCR assay for rapid detection of herpes simplex virus in cerebrospinal fluid. *J Clin Microbiol* 38:3214–3218. <https://doi.org/10.1128/JCM.38.9.3214-3218.2000>
- Szpara ML, Petrussevska RT, Breikreutz D, Hornung J, Markham A, Fusenig NE. 1988. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *J Cell Biol* 106:761–771. <https://doi.org/10.1083/jcb.106.3.761>
- Szpara ML, Tafuri YR, Enquist LW. 2011. Preparation of viral DNA from nucleocapsids. *J Vis Exp* 54:2–7. <https://doi.org/10.3791/3151>
- Szpara M. 2014. Isolation of herpes simplex virus nucleocapsid DNA, p 31–41. In Diefenbach RJ, Fraefel C (ed), *Herpes simplex virus*. Springer, New York, New York, NY.
- Parsons LR, Tafuri YR, Shreve JT, Bowen CD, Shipley MM, Enquist LW, Szpara ML. 2015. Rapid genome assembly and comparison decode intrastrain variation in human alphaherpesviruses. *MBio* 6:e02213-14. <https://doi.org/10.1128/mBio.02213-14>
- Hannon GJ. 2010. FASTX-toolkit FASTQA short-reads pre-process tools. FASTX-toolkit. http://hannonlab.cshl.edu/fastx_toolkit
- 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>
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>
- Warren RL, Sutton GG, Jones SJM, Holt RA. 2007. Assembling millions of short DNA sequences using SSAKE. *Bioinformatics* 23:500–501. <https://doi.org/10.1093/bioinformatics/btl629>
- Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, et al. 2000. A whole-genome assembly of *Drosophila*. *Science* 287:2196–2204. <https://doi.org/10.1126/science.287.5461.2196>
- Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol* 13:1–9. <https://doi.org/10.1186/gb-2012-13-6-r56>
- Angiuoli SV, Salzberg SL. 2011. Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics* 27:334–342. <https://doi.org/10.1093/bioinformatics/btq665>
- Dolan A, Jamieson FE, Cunningham C, Barnett BC, McGeoch DJ. 1998. The genome sequence of herpes simplex virus type 2. *J Virol* 72:2010–2021. <https://doi.org/10.1128/JVI.72.3.2010-2021.1998>
- Forni D, Pontremoli C, Clerici M, Pozzoli U, Cagliani R, Sironi M. 2020. Recent out-of-Africa migration of human herpes simplex viruses. *Mol Biol Evol* 37:1259–1271. <https://doi.org/10.1093/molbev/msaa001>
- Chang W, Jiao X, Sui H, Goswami S, Sherman BT, Fromont C, Caravaca JM, Tran B, Imamichi T. 2022. Complete genome sequence of herpes simplex virus 2 strain G. *Viruses* 14:536. <https://doi.org/10.3390/v14030536>
- López-Muñoz AD, Rastrojo A, Kropp KA, Viejo-Borbolla A, Alcamí A. 2021. Combination of long- and short-read sequencing fully resolves complex repeats of herpes simplex virus 2 strain MS complete genome. *Microb Genom* 7. <https://doi.org/10.1099/mgen.0.000586>
- Pyöriä L, Jokinen M, Toppinen M, Salminen H, Vuorinen T, Hukkanen V, Schmotz C, Elbasani E, Ojala PM, Hedman K, Välimaa H, Perdomo MF. 2020. HERQ-9 is a new multiplex PCR for differentiation and quantification of all nine human herpesviruses. *mSphere* 5:e00265-20. <https://doi.org/10.1128/mSphere.00265-20>
- Paavilainen H. 2017. Inhibition of herpes simplex virus infection with RNA interference PhD dissertation, University of Turku, Turku, Finland
- Colgrove R, Diaz F, Newman R, Saif S, Shea T, Young S, Henn M, Knipe DM. 2014. Genomic sequences of a low passage herpes simplex virus 2 clinical isolate and its plaque-purified derivative strain. *Virology (Auckl)* 450–451:140–145. <https://doi.org/10.1016/j.virol.2013.12.014>
- López-Muñoz AD, Rastrojo A, Alcamí A. 2018. Complete genome sequence of herpes simplex virus 2 strain 333. *Microbiol Resour Announc* 7:e00870-18. <https://doi.org/10.1128/MRA.00870-18>
- Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166. <https://doi.org/10.1093/bib/bbx108>