Genome Sequence of the Virulent Model Herpes Simplex Virus 1 Strain McKrae Demonstrates the Presence of at Least Two Widely Used Variant Strains

Daniel W. Renner,[a,c] Lance Parsons,[b] Jacob T. Shreve,[c] Esteban A. Engel,[d] Chad V. Kuny,[a,c] Lynn Enquist,[e] Donna Neumann,[f] Colleen Mangold,[a,g] Moriah L. Szpara[a,c]

[a]Department of Biology, Pennsylvania State University, University Park, Pennsylvania, USA
[b]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey, USA
[c]Department of Biochemistry and Molecular Biology, Center for Infectious Disease Dynamics, Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, Pennsylvania, USA
[d]Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, USA
[e]Department of Molecular Biology, Princeton University, Princeton, New Jersey, USA
[f]Department of Ophthalmology and Visual Sciences, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA
[g]Department of Entomology, Pennsylvania State University, University Park, Pennsylvania, USA

ABSTRACT Herpes simplex virus 1 (HSV-1) strain McKrae was isolated in 1965 and has been utilized by many laboratories. Three HSV-1 strain McKrae stocks have been sequenced previously, revealing discrepancies in key genes. We sequenced the genome of HSV-1 strain McKrae from the laboratory of James M. Hill to better understand the genetic differences between isolates.

Herpes simplex virus 1 (HSV-1, or human herpesvirus 1, HHV-1) is a neurotropic and neuroinvasive double-stranded DNA virus of the *Herpesviridae* family (genus *Simplexvirus*, subfamily *Alphaherpesvirinae*), which infects humans. HSV-1 strains have been shown experimentally to exhibit wide variation in both virulence and pathogenic phenotypes (1, 2). The HSV-1 strain McKrae was originally isolated from a human sample and has been passaged through multiple cell types and species (3). McKrae is one of the most highly virulent model strains of HSV-1, demonstrating a high reactivation frequency and neurovirulence in animal models (1, 4–6). There are currently three variant strains of McKrae with genome sequences in GenBank (submitted under accession no. JX142173.1 [7], JQ730035.1 [8], and MN136524 [9]). The differences between these published genome sequences—for instance, the putative JQ730035.1 truncations of the UL56 and UL36 open reading frames (ORFs) and an extension of the US10 ORF (see reference 8 for details)—led us to sequence the genome of HSV-1 McKrae from the laboratory of James M. Hill (located at Louisiana State University Health Sciences Center until his death in 2013). Hill often utilized stocks of HSV-1 McKrae in his comparisons of HSV-1 virulence in mice and rabbits, and he freely shared these resources with other labs (1, 10, 11).

We isolated viral nucleocapsid DNA from HSV-1 McKrae-infected Vero (African green monkey kidney) cells, using published methods (12, 13). Cells infected at a high multiplicity of infection (MOI = 10) were scraped, host membranes were disrupted via Freon extraction, and viral capsids were separated from cellular debris by pelleting through a density gradient. After detergent and proteinase K lysis of capsids, the viral nucleocapsid DNA was further purified using phenol-chloroform extraction, followed by ethanol precipitation of the viral genomic DNA. We followed the manufacturer's protocols (Illumina TruSeq DNA) to produce a barcoded library of 500-bp fragments and to obtain 16.3 million 100-bp, paired-end sequence reads (Illumina HiSeq 2000 with v2 chemistry).

The viral genome assembly process has been previously described (14) and is

FIG 1 Strain McKrae genomes are closely related but do not all express intact UL56 protein. (A) A UPGMA tree built based on a whole-genome alignment of four strain McKrae genome sequences demonstrates their high relatedness (excluding gaps). (B) However, the UL56 gene in the Hill lab McKrae and two other published McKrae genome sequences has a frameshift and early truncation, which results in a lack of UL56 protein. Western blots of protein lysates are shown from epithelial Vero (monkey kidney) cells at 6 hours postinfection (hpi), including a mock infection, as well as HSV-1 strains F, McKrae (Hill lab; GenBank accession no. KT425107), and KOS. Viral glycoprotein E (gE) demonstrates infection for all three strains, and beta-actin serves as a loading control. The conditions for Vero cell infection and Western blotting were the same as those for our previously published studies (21, 22). Antibody sources: mouse anti-beta-actin, Sigma-Aldrich; mouse anti-gE, kindly provided by Harvey Friedman; rabbit anti-UL56 antibody, generated by Esteban Engel.

summarized here, with full installation and usage instructions at http://virga.readthedocs .org/. All scripts were used with default parameters unless stated otherwise. For quality control prior to genome assembly, we first trimmed any library adapter sequences, sequencing artifacts, and low-quality bases using FastX-Toolkit and Trimmomatic (15). Then, Bowtie 2 was used to map and remove any host-derived DNA sequences (using the rhesus macaque genome sequence as a proxy for Vero cell DNA) (14, 16). After quality control and host-contaminant removal, 11.7 million paired-end sequence reads were used as the input for multiple SSAKE v3.8 *de novo* assemblies (using eight pairs of overlap and trimming parameters, 16-0, 19-0, 20-0, 29-0, 16-4, 19-4, 20-4, and 29-4). These SSAKE contigs were combined into larger sequence blocks using Celera v8.1 (14). We then used Mugsy v2.3 to align these contigs to the HSV-1 reference genome sequence (GenBank accession no. JN555585) and stitched them into a consensus genome sequence using a custom script, maf_net.py (https:// bitbucket.org/szparalab/virga/src/master/), and GapFiller v1.10 (insert_deviation 0.2) (14). Finally, we used a genome-wide alignment (ClustalW2 v2.1) of the new consensus genome sequence with the HSV-1 reference genome sequence (JN555585) as a template to transfer annotations, using the custom script compare_genomes.py (https://bitbucket.org/szparalab/virga/src/master/) (14). This new HSV-1 McKrae genome sequence (KT425107) is 149,310 bp, with 6 gaps in comparison to the HSV-1 reference genome sequence (JN555585). The new McKrae genome sequence has 67.8% GC content and an average sequence depth of coverage exceeding 1,000× (with 97.5% of the genome at ≥100× depth).

We performed genome alignments between this McKrae genome sequence (KT425107) and those previously published by the Jin laboratory (JQ730035) (8), the Morrison and Davido laboratories (JX142173) (7), and the Imamichi laboratory (MN136524) (9), using MAFFT v7.475 (FFT-NS-2 with default parameters; https://mafft .cbrc.jp/alignment/server/). We generated a tree from all gap-free sites in the full-genome sequence alignment (145,746 bases), using the UPGMA method (Jukes-Cantor model; Geneious v11.1.5). This indicated that three of the McKrae genomes are very closely related, with JX142173 from the Morrison and Davido laboratories separated by

a small margin (Fig. 1A). With the exception of the UL56 variant noted below, none of the ORF truncations or extensions described in the prior publication on McKrae strain JQ730035 (8) were detected in this new McKrae genome.

One of the major differences between the prior published variants of HSV-1 McKrae was in the UL56 gene—it has a truncated open reading frame (ORF) in JQ730035 and MN136524 and an intact ORF similar to other HSV-1 strains in JX142173 (17). We found an identical truncation in the Hill lab McKrae strain, KT425107 (the current genome sequence), and the previously described McKrae strains JQ730035 and MN136524. In contrast, the UL56 ORF in the Morrison and Davido laboratories' McKrae strain JX142173 is full length and appears to be intact (7). The nonintact UL56 genes have a frameshift (due to an extra G nucleotide) after amino acid 97, an altered amino acid tail, and an early stop codon at amino acid 181 (versus a full-length protein of 234 amino acids in other strains). We have verified the absence of the UL56 protein product in the Hill lab McKrae strain KT425107 (Fig. 1B). While an intact UL56 gene has been shown to be required for HSV-1 virulence *in vivo* (18–20), all strains used in those prior studies are less virulent in animal models than HSV-1 McKrae (1, 2, 10, 11). These data raise the question of whether or not the UL56 protein is required in the HSV-1 McKrae genetic background to achieve full virulence in animal models. We encourage other researchers to evaluate the status of the UL56 gene product in their own stocks of this model virulent HSV-1 strain to help address this question.

Data availability. The HSV-1 McKrae genome sequence has been deposited at GenBank under the accession no. KT425107. The raw reads have been submitted to the SRA as accession no. SRR13801763.

ACKNOWLEDGMENTS

REFERENCES

1. Hill JM, Rayfield MA, Haruta Y. 1987. Strain specificity of spontaneous and adrenergically induced HSV-1 ocular reactivation in latently infected rabbits. Curr Eye Res 6:91–97. https://doi.org/10.3109/02713688709020074.
2. Dix RD, McKendall RR, Baringer JR. 1983. Comparative neurovirulence of herpes simplex virus type 1 strains after peripheral or intracerebral inoculation of BALB/c mice. Infect Immun 40:103–112. https://doi.org/10.1128/IAI.40.1.103-112.1983.
3. Williams LE, Nesburn AB, Kaufman HE. 1965. Experimental induction of disciform keratitis. Arch Ophthalmol 73:112–114. https://doi.org/10.1001/archopht.1965.00970030114023.
4. Sawtell NM, Poon DK, Tansky CS, Thompson RL. 1998. The latent herpes simplex virus type 1 genome copy number in individual neurons is virus strain specific and correlates with reactivation. J Virol 72:5343–5350. https://doi.org/10.1128/JVI.72.7.5343-5350.1998.
5. Yao H-W, Ling P, Chen S-H, Tung Y-Y, Chen S-H. 2012. Factors affecting herpes simplex virus reactivation from the explanted mouse brain. Virology 433:116–123. https://doi.org/10.1016/j.virol.2012.07.018.
6. Wang H, Davido DJ, Morrison LA. 2013. HSV-1 strain McKrae is more neuroinvasive than HSV-1 KOS after corneal or vaginal inoculation in mice. Virus Res 173:436–440. https://doi.org/10.1016/j.virusres.2013.01.001.
7. Macdonald SJ, Mostafa HH, Morrison LA, Davido DJ. 2012. Genome sequence of herpes simplex virus 1 strain McKrae. J Virol 86:9540–9541. https://doi.org/10.1128/JVI.01469-12.
8. Watson G, Xu W, Reed A, Babra B, Putman T, Wick E, Wechsler SL, Rohrmann GF, Jin L. 2012. Sequence and comparative analysis of the genome of HSV-1 strain McKrae. Virology 433:528–537. https://doi.org/10.1016/j.virol.2012.08.043.
9. Jiao X, Sui H, Lyons C, Tran B, Sherman BT, Imamichi T. 2019. Complete genome sequence of herpes simplex virus 1 strain McKrae. Microbiol Resour Announc 8:e00993-19. https://doi.org/10.1128/MRA.00993-19.
10. Al-Dujaili LJ, Clerkin PP, Clement C, McFerrin HE, Bhattacharjee PS, Varnell ED, Kaufman HE, Hill JM. 2011. Ocular herpes simplex virus: how are latency, reactivation, recurrent disease and therapy interrelated? Future Microbiol 6:877–907. https://doi.org/10.2217/fmb.11.73.
11. Webre JM, Hill JM, Nolan NM, Clement C, McFerrin HE, Bhattacharjee PS, Hsia V, Neumann DM, Foster TP, Lukiw WJ, Thompson HW. 2012. Rabbit and mouse models of HSV-1 latency, reactivation, and recurrent eye diseases. J Biomed Biotechnol 2012:612316. https://doi.org/10.1155/2012/612316.
12. Szpara ML, Tafuri YR, Enquist LW. 2011. Preparation of viral DNA from nucleocapsids. J Vis Exp e3151. https://doi.org/10.3791/3151.
13. 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, NY.
14. Parsons LR, Tafuri YR, Shreve JT, Bowen CD, Shipley MM, Enquist LW, Szpara ML. 2015. Rapid genome assembly and comparison decode intra-strain variation in human alphaherpesviruses. mBio 6:e02213-14. https://doi.org/10.1128/mBio.02213-14.
15. 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.

16. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

17. Szpara ML, Gatherer D, Ochoa A, Greenbaum B, Dolan A, Bowden RJ, Enquist LW, Legendre M, Davison AJ. 2014. Evolution and diversity in human herpes simplex virus genomes. J Virol 88:1209–1227. https://doi.org/10.1128/JVI.01987-13.

18. Berkowitz C, Moyal M, Rösen-Wolff A, Darai G, Becker Y. 1994. Herpes simplex virus type 1 (HSV-1) UL56 gene is involved in viral intraperitoneal pathogenicity to immunocompetent mice. Arch Virol 134:73–83. https://doi.org/10.1007/BF01379108.

19. Kehm R, Rosen-Wolff A, Darai G. 1996. Restitution of the UL56 gene expression of HSV-1 HFEM led to restoration of virulent phenotype; deletion of the amino acids 217 to 234 of the UL56 protein abrogates the virulent phenotype. Virus Res 40:17–31. https://doi.org/10.1016/0168-1702(96)80248-6.

20. Kolb AW, Lee K, Larsen I, Craven M, Brandt CR. 2016. Quantitative trait locus based virulence determinant mapping of the HSV-1 genome in murine ocular infection: genes involved in viral regulatory and innate immune networks contribute to virulence. PLoS Pathog 12:e1005499. https://doi.org/10.1371/journal.ppat.1005499.

21. Shipley MM, Renner DW, Ott M, Bloom DC, Koelle DM, Johnston C, Szpara ML. 2018. Genome-wide surveillance of genital herpes simplex virus type 1 from multiple anatomic sites over time. J Infect Dis 218:595–605. https://doi.org/10.1093/infdis/jiy216.

22. Mangold CA, Rathbun MM, Renner DW, Kuny CV, Szpara ML. Viral infection of human neurons triggers strain-specific differences in host neuronal and viral transcriptomes. PLoS Pathog, in press.