


SCIENTIFIC REPORTS



OPEN

Inferred father-to-son transmission of herpes simplex virus results in near-perfect preservation of viral genome identity and *in vivo* phenotypes

Utsav Pandey¹, Daniel W. Renner¹, Richard L. Thompson², Moriah L. Szpara¹ ¹ & Nancy M. Sawtell³

High throughput sequencing has provided an unprecedented view of the circulating diversity of all classes of human herpesviruses. For herpes simplex virus 1 (HSV-1), we and others have previously published data demonstrating sequence diversity between hosts. However the extent of variation during transmission events, or in one host over years of chronic infection, remain unknown. Here we present an initial example of full characterization of viruses isolated from a father to son transmission event. The likely occasion of transmission occurred 17 years before the strains were isolated, enabling a first view of the degree of virus conservation after decades of recurrences, including transmission and adaptation to a new host. We have characterized the pathogenicity of these strains in a mouse ocular model of infection, and sequenced the full viral genomes. Surprisingly, we find that these two viruses have preserved their phenotype and genotype nearly perfectly during inferred transmission from father to son, and during nearly two decades of episodes of recurrent disease in each human host. Given the close genetic relationship of these two hosts, it remains to be seen whether or not this conservation of sequence will occur during non-familial transmission events.

Herpes simplex viruses (HSV) are widespread human pathogens. HSV-1 and HSV-2 together have ~90% incidence worldwide¹. HSV infections are a major public health concern, causing mucocutaneous and systemic diseases². In the USA, HSV-1 is the leading cause of sporadic necrotizing encephalitis and infectious blindness^{3,4}. HSV infection remains a significant cause of morbidity and mortality in neonates, with an incidence of approximately 1 per 3,200 deliveries in the United States^{5,6}. HSV-2, which has 70% genomic similarity to HSV-1, is associated with higher HIV/AIDS acquisition in developing countries⁷⁻⁹. Traditionally, epidemiologically related strains of HSV-1 or HSV-2 have been identified by comparing restriction fragment length polymorphism (RFLP) patterns or using targeted Sanger sequencing of selected genes or loci¹⁰⁻¹³. Although RFLPs were crucial to establish our overall understanding of HSV transmission, the limited number of nucleotides assayed by changes in RFLP pattern obscured variation at the level of single nucleotide variations (SNVs) and small insertions or deletions (in/dels). Advancements in high-throughput sequencing (HTS) have now made it possible to study the full extent of genetic variation in the viral population being transmitted between different hosts.

A transmission event has the potential to create a bottleneck that reduces genetic variation, if only a limited number of founder viruses are transmitted to the next host¹⁴⁻¹⁸. Viral genetic diversity may be created *de novo* through replication, recombination, and/or selection in the new host, leading to a viral population that is genetically distinct from the founder population. Most transmission studies have focused on RNA viruses, which have a larger amount of standing variation in the population^{19,20}. Less is known about bottlenecks or expansion

¹Department of Biochemistry and Molecular Biology, Center for Infectious Disease Dynamics, and the Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, Pennsylvania, 16802, USA. ²Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati, Cincinnati, Ohio, 45229, USA. ³Division of Infectious Diseases, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, 45229, USA. Correspondence and requests for materials should be addressed to M.L.S. (email: moriah@psu.edu)

of variation during transmission of large DNA viruses. The transmission bottleneck has been shown to be more restrictive for RNA viruses than for DNA viruses^{17,18}. In the case of human immunodeficiency virus (HIV) and hepatitis C virus (HCV), it is estimated that as few as 1–5 virions may be transmitted between hosts to establish infection^{14–16}. In contrast, the founder population for human cytomegalovirus (HCMV), a β -herpesvirus, was estimated to be in the tens to hundreds of virions for maternal-to-fetal transmission¹⁷. While the conservation of genetic diversity is thought to be more likely during transmission of DNA viruses, data on viral genetic diversity during human-to-human transmission of HSV-1 has thus far been lacking.

Here we present the genetic and phenotypic characterization of HSV-1 isolates obtained from a father and his son. These isolates were obtained from oral lesions of each host, nearly two decades after the inferred transmission and primary infection of the son. We investigated the pathogenicity of these isolates in a mouse ocular model of infection, and found a similar level of replication at the site of inoculation, and a similar level of neuroinvasiveness. The viral isolates were also alike in their ability to establish latency and their rate of reactivation in both *in vivo* and *in vitro* models. We subjected the viral population of each cultured viral isolate to HTS, and found that the genomes were nearly identical at the consensus level. A small number of minority variants were found in each isolate, most of which result from length variations in homopolymeric tracts or short sequence repeats (SSRs). In the light of recent studies showing extensive intra-host variation and rapid evolution during transmission of HCMV^{17,21}, the extent of conservation we observed between the isolates was exceptional. This study presents the first example of comprehensive characterization of founder and transmitted HSV-1 isolates using HTS and *in vivo* models of pathogenesis.

Results

Familial transmission and viral culture characteristics. The likely timing of horizontal transmission from father to son, and the collection of the virus isolates from each, is shown schematically in Fig. 1. Both the father and the son experience recurrences 6–8 times per year. The father harbored the virus for 24 years before the likely transmission to the son. The timing of transmission is inferred from the son's primary episode of gingivostomatitis. The mother was seronegative, and the child had no contact with other caregivers or siblings before this time. Virus was isolated from a recurrence in the father 43 years after his primary infection (isolate R-13), and from a recurrence in the son 17 years after the primary infection. The isolates from father and son were collected on independent occasions, two years apart. No differences in plaque size or morphology were noted between the isolates upon culturing. Both virus isolates were found to replicate with equivalent efficiency. RFLPs were examined in a series of Southern blots on DNA cleaved with several different restriction enzymes, utilizing diverse simple probes (~1 kb) or complex cosmid probes of 27 to 45 kb in length that spanned the entire viral genome (see Supplementary Fig. 1 for a representative example). At this level of resolution, no genetic differences were evident between the two isolates.

Acute replication kinetics and mortality of HSV-1 isolates R-13 and N-7. The *in vivo* pathogenesis properties of the isolates were examined in the mouse ocular model of infection. Groups of outbred male Swiss Webster mice were infected via the cornea with 2×10^5 PFU of father and son isolates R-13 and N-7, the unrelated clinical isolate CI-37, or the HSV-1 reference strain 17syn+ (see Methods for details). At two day intervals over a 10 day period, tissues were harvested from three mice in each group and the virus content quantified by standard plaque titration assays. Replication kinetic curves are shown in Fig. 2A–C. There was no significant difference in peripheral replication on the eye between the father-son viral isolates or HSV-1 17syn+, however the viral titer for CI-37 was significantly higher (Fig. 2A) ($p \leq 0.05$; see Fig. 2 for summed value of each area under the curve (AUC)). Likewise, virus replication in trigeminal ganglia (TG) was not significantly different between R-13, N-7 and 17syn+, but was significantly higher for isolate CI-37 (Fig. 2B) ($p \leq 0.01$; see Fig. 2 for AUC values). In the ocular model of infection, the central nervous system (CNS) is infected by retrograde transport of the virus from infected neurons in the TG to the trigeminal nucleus in the hindbrain. From there, replicating virus is transported toward the front of the brain through time. To capture some of this process, brain tissue was divided into four roughly equal parts for assay (rear, mid rear, mid front, front) as detailed previously²². No significant differences were observed between the father and son viral isolates or HSV-1 17syn+ in any of the brain regions, whereas CI-37 had a significantly higher titer in all four brain regions (Fig. 2C) ($p \leq 0.01$ for rear and mid rear regions, $p \leq 0.05$ for mid front and front regions; see Fig. 2 for AUC values). The amount of virus detected in the brain for R-13 and N-7 was low, reaching just 100 PFU in the rear portion of the brain and less in the front portions. In contrast, the viral isolate CI-37 reached much higher titers of 10^4 to 10^5 PFU in each brain region, while 17syn+ had an intermediate phenotype. These data were consistent with the 100% survival observed with both R-13 ($n = 35$) and N-7 ($n = 34$), the complete mortality of CI-37 ($n = 5$), and the 81% survival of 17syn+ ($n = 16$) (Supplementary Fig. 2). No signs of encephalitis in either R-13 or N-7 infected groups were observed, although moderate to severe blepharitis did occur.

Since survival rates were high for the father-son isolates R-13 and N-7, additional mice were maintained for 40 days post-infection (p.i.) and TG and brain tissues were analyzed for the presence of the latent viral genomes by quantitative PCR (see Methods for details). There was no significant difference in the number of viral genomes detected in mouse TG that were latently infected with either R-13 or N-7 (Fig. 2D). Likewise, no significant difference in the amount of latent viral DNA present in the four regions of the brain was detected (Fig. 2E). These results were consistent with the prior observation that the isolates replicated equivalently in these tissues.

The remaining latently-infected mice were employed to test the ability of the isolates to reactivate from latency. Reactivation was tested after explantation of TG into culture *in vitro*²³ or after hyperthermic stress (HS) *in vivo*^{24–26}. No infectious virus was detected in any of six TG tested at 0 hours post-explant in either group, demonstrating that persistent infection or frequent spontaneous reactivation events were absent for both R-13 and N-7²⁷ (Fig. 3A and D). After three days in culture all TG samples were positive for virus (Fig. 3A). We conclude from this that

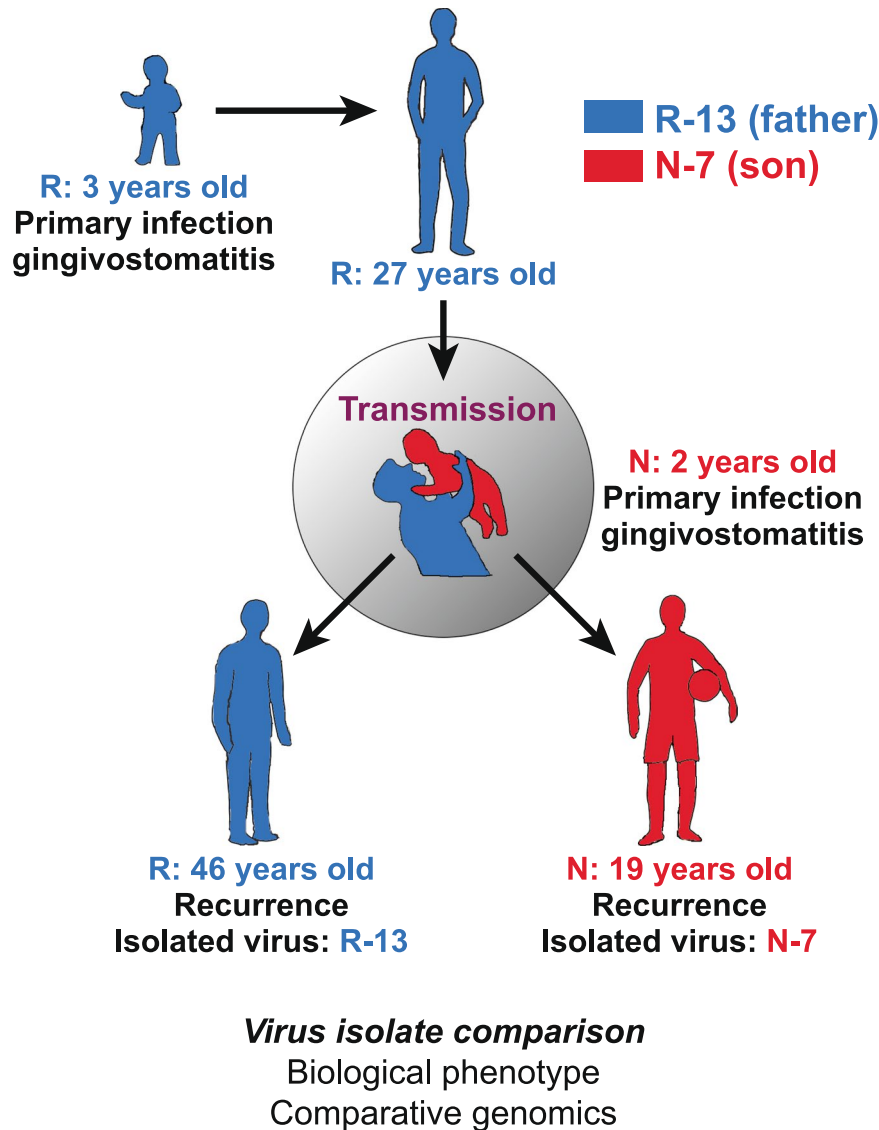


Figure 1. Timing of father to son transmission of HSV and viral isolate acquisition. HSV-1 isolate R-13 was acquired during an orolabial recurrence from a father with lifelong chronic infection, characterized by oral lesion recurrences 6–8 times per year. The father acquired his primary infection at age 3, more than four decades prior. At age 27, the father is inferred to have transmitted HSV-1 to his 2-year old son, who experienced gingivostomatitis with high fever. After resolution of the primary infection, the son likewise experienced recurrent orolabial lesions at a frequency of 6–8 episodes/year. Isolate N-7 was acquired during an orolabial recurrence from the son 17 years after his primary infection. The isolates were obtained on separate occasions, two years apart.

latent virus capable of reactivating was present in all TG tested. Whole ganglion immunohistochemical (IHC)^{28,29} detection of viral proteins revealed regions of virus spread within the TG at this time (Fig. 3E). Two separate experiments showed similar efficiency of viral reactivation *in vivo* in TG 22 hrs. post HS, with 50% of the TG examined in both groups containing detectable infectious virus (Fig. 3B). A similar number of TG neurons were found to be expressing viral proteins at 22 hrs. post HS, as detected by whole ganglion IHC (Fig. 3C and F,G). Combined with the finding above that similar numbers of latent viral genomes were present in these TG, we concluded that the relative efficiency of reactivation from latency *in vivo* was also similar between the father-son isolates R-13 and N-7.

Nearly identical genomes of father and son HSV-1 isolates. To further test the observation of overall genomic and phenotypic similarity of the HSV-1 isolates from father (R-13) and son (N-7), we used HTS and *de novo* assembly approaches to construct a full-length genome sequence for each viral isolate. Viral nucleocapsid DNA was used as the starting point for Illumina paired-end sequencing, with *de novo* assembly and genome annotation carried out using a recently described open-source viral genome assembly (VirGA) workflow³⁰ (Table 1) (see Methods for details). At the level of consensus genomes, HSV-1 isolates R-13 and N-7 are 98.9%

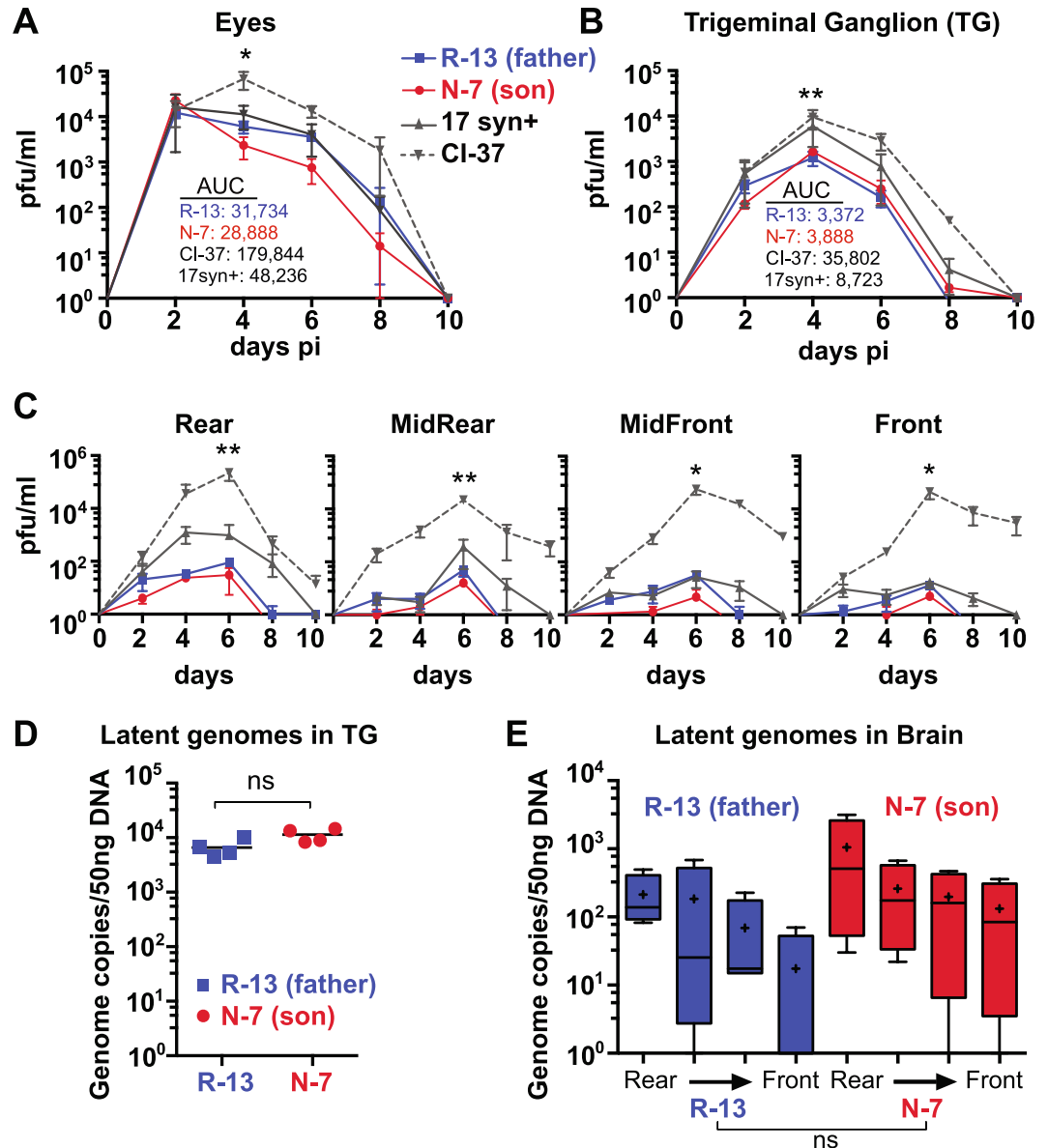


Figure 2. Quantification of replication and latency phenotypes of R-13 and N-7 during infection *in vivo*. Swiss Webster mice (male) were infected on scarified corneas with 2×10^5 pfu of the clinical isolate R-13 or N-7. At the indicated times post-infection, tissues collected from each of three mice per group were assayed for infectious virus using a standard plaque assay (see Methods for details). Replication in eyes (A) and TG (B) revealed that infectious virus generated in these tissues during the acute stage of infection was not significantly different between R-13, N-7, or 17syn+, whereas the levels for CI-37 were significantly higher (Student's t test, $*p \leq 0.05$, $**p \leq 0.01$, for peak titer on day 4; AUC = area under the curve) (C) The replication kinetics and regional distribution of isolates in the brain was determined by cutting each brain into 4 coronal sections. The levels of infectious virus were not significantly different between R-13, N-7, or 17syn+, whereas the levels for CI-37 were significantly higher (Student's t test, $*p \leq 0.05$, $**p \leq 0.01$, for peak titer on day 6). This low level of infectious virus for R-13 and N-7 in the brain (≤ 100 PFU) is consistent with the 100% survival observed for both isolates under these infection conditions, whereas CI-37 induced complete mortality (Supplementary Fig. 2). (D,E) Quantification of latent viral genomes. At 40 days post infection, the TG and brains from R-13 ($n = 4$) and N-7 ($n = 4$) latently infected mice were assayed for viral genome copies using real time qPCR. Viral genome copy numbers per 50 ng mouse DNA detected in R-13 and N-7 were not significantly different (Student's t test) (D). Brains were cut into four coronal sections prior to isolating DNA and the number of viral genome copies was determined in each section. Viral genome copies in the brains were not significantly different (ANOVA; on box and whisker plot, + indicates mean, bar indicates median) (E).

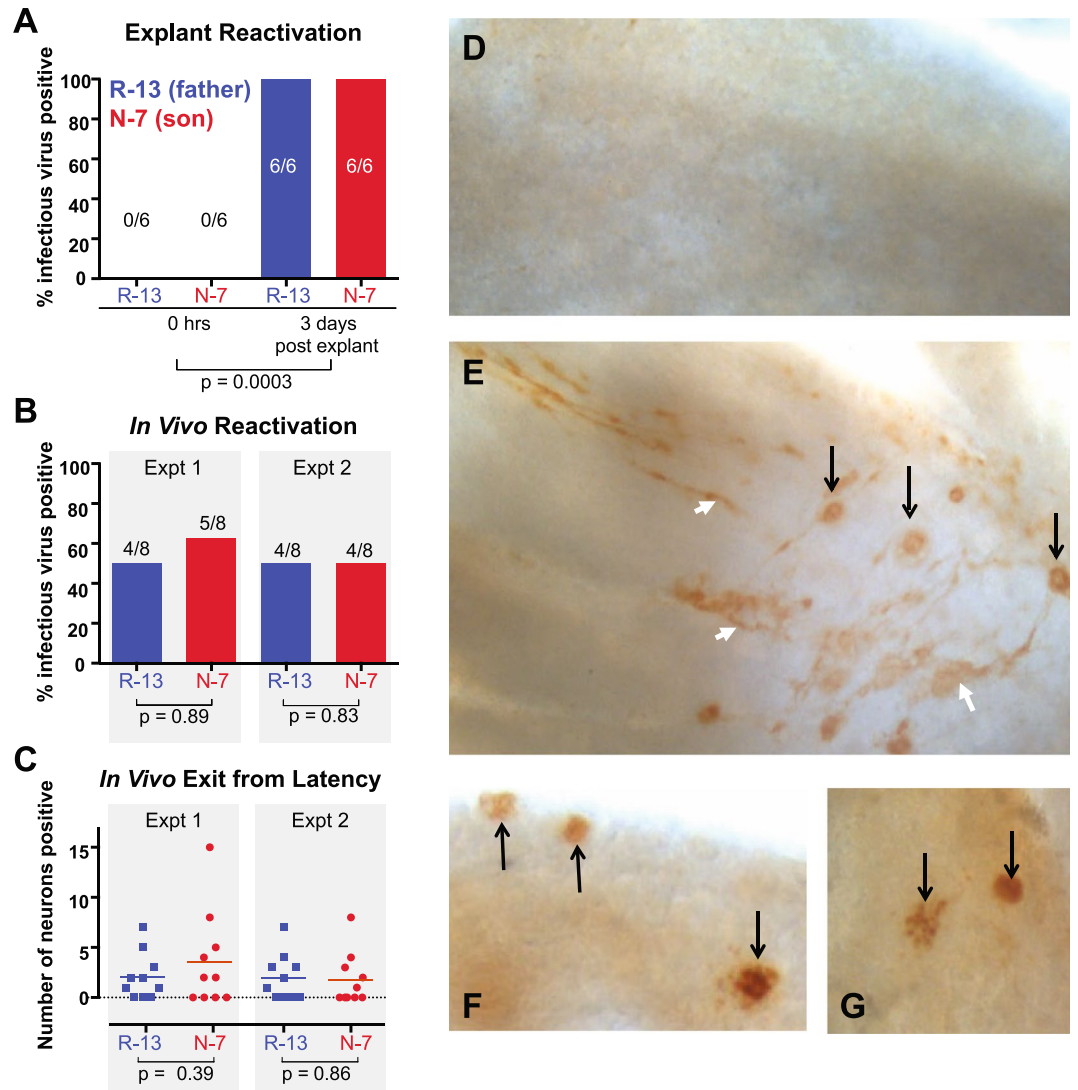


Figure 3. Explant and *In vivo* reactivation in Swiss Webster mice latently infected with R-13 and N-7. R-13 and N-7 were compared for reactivation from latency (>40 days pi) using *in vitro* and *in vivo* reactivation assays. (A) In a standard TG explant assay, no difference between R-13 and N-7 reactivation frequency was observed ($p = 0.99$, Student's t test) although the difference between virus recovered at time 0 and 3 days post explant was significant in both groups ($p = 0.0003$, ANOVA). (B) The *in vivo* reactivation frequency (percentage of mice with infectious virus detected in TG) was also not different between R-13 and N-7 at 22 hrs. post hyperthermic stress (Student's t test, expt. 1 $p = 0.89$, expt. 2 $p = 0.83$) and (C) the number of neurons exiting latency was also not different (expt. 1 $p = 0.39$, expt. 2 $p = 0.86$). Latently infected TG were subjected to whole ganglion immunohistochemistry to detect viral protein at 0 hrs. (D), and 3 days (E) post explant. Viral protein-expressing neurons (black arrows) and tracts (white arrowheads) mark the range of viral spread. Viral protein-expressing neurons (black arrows) are detectable in TG from N-7 (F) and R-13 (G), which reactivated from latently infected TG, at 22 hrs. after hyperthermic stress *in vivo*.

HSV-1 strain	Paired-end read length	Raw reads	Used for assembly	Genome length	Average depth (X-fold)	GenBank Accession
R-13	300 bp	3.1×10^6	2.4×10^6	151,636	7,626	KY922718
N-7	300 bp	4.3×10^6	3.6×10^6	151,765	11,640	KY922719

Table 1. Sequencing statistics for R-13 and N-7 strains of HSV-1.

identical in their DNA sequence (Table 2). This level of identity between samples has previously only been seen in the case of sub-clones picked or isolated from a common parental virus stock^{30–32}.

For the 1.1% of the genome that did differ between HSV-1 isolates R-13 and N-7, we categorized these differences as genic (coding) vs. intergenic, and grouped insertions or deletions (in/dels) vs. single-nucleotide

Comparisons between strains	DNA identity	Intergenic		Genic		
		in/dels (# events)	SNVs	in/dels (# events)	Synonymous SNVs	Non-synonymous SNVs
N-7 vs. R-13	98.9%	16	2	2*	0	0

Table 2. Pair-wise DNA identity and variant proteins between HSV-1 consensus genomes. *54 bp insertion in VP1/2 (UL36) and a 6 bp insertion in ICP4 (RS1) in N-7, relative to R-13.

variations (SNVs) (Table 2). We found no SNVs in any coding sequence, and only two in intergenic regions. For the in/dels, we calculated the minimum number of insertion or deletions events that could have led to the observed differences. For instance, a three base pair insertion was counted as one in/del event. There were a total of 18 in/del events, of which only two occurred in coding regions. These included an 54 base pair (18 amino acid (AA)) in/del in the gene encoding the ubiquitin-specific protease, VP1/2 (UL36), and a 6 base pair (2 AA) in/del in the transcriptional regulator protein, ICP4 (RS1) (Table 2). The insertion in VP1/2 is present in the C-terminal region of the protein, which contains an extended array of ‘PQ’ tandem repeats. Sequence length fluctuation in this region has been documented in HSV and other alphaherpesviruses such as varicella zoster virus (VZV), Marek’s disease virus (MDV), and pseudorabies virus (PRV)^{33–37}. Similarly, the insertion observed in ICP4 is also present in a short sequence repeat in a functionally uncharacterized domain of the protein³⁸. In both cases, the N-7 viral genome from the son contains the longer sequence (i.e. has an insertion) relative to the R-13 (father’s) genome.

Comparison of father and son isolates to other HSV-1 strains. To place the HSV-1 isolates from father (R-13) and son (N-7) in the context of previously sequenced HSV-1 isolates, we compared full-length genomes of R-13 and N-7 to all available HSV-1 genomes in GenBank (46 total; see Methods for full list). We investigated the relatedness of the genomes by constructing a network graph using SplitsTree (Supplementary Fig. S3). We observed that the father-son isolates R-13 and N-7 form a separate branch compared to all previously sequenced genomes, but are positioned in the tree graph between other Asian, European and African isolates. The phylogenetic separation of isolates based on geography has been previously described³¹.

To put the protein-coding features of HSV-1 isolates R-13 and N-7 in the context of prior work, we compared all protein (AA) sequences of these isolates to those encoded by 46 fully-sequenced HSV-1 strains. In this comparison, we found a total of 34 viral proteins containing AA variations in R-13 and N-7 that had never been observed before (see Supplementary Table 1 for a complete list). This large number of unique residues may reflect the position of these isolates in the phylogenetic network, where they lie on a branch separate from other known strains. We also found 54 proteins with AA variations that differ from the HSV-1 reference genome (strain 17syn+, Genbank ID JN555585), but that have been observed previously in one or more of the other sequenced strains.

Intra-strain variation: Detection of minority variants within each HSV-1 genome. Comparison of the consensus genomes of the HSV-1 isolates from father (R-13) and son (N-7) provided a global view of the overall identity of these viral isolates. However as observed for HCMV and VZV^{17,21,39}, intra-strain variation can exist within a viral population, or within a local niche or lesion of a given host. To assess the extent of genetic diversity within each viral isolate, we investigated whether the consensus genome of either R-13 or N-7 contained any minority variants, or nucleotide positions where HTS data indicated more than one possible allele. This allowed us to detect specific nucleotide bases (loci) for which another allele was present; these minor differences would otherwise be missed in a comparison of the majority-genotype (consensus level genome) as seen in Table 2. We observed 59 minority variants in the R-13 genome and 48 in the N-7 genome (Supplementary Table 2). Most of the minor alleles observed in R-7 and N-13 genomes occurred at a low frequency or/and were intergenic. These minority variants were neutral and unlikely to have a major effect on viral fitness. These results are consistent with recent findings in HCMV, where it was shown that the vast majority of these minority variants observed were neutral⁴⁰. Although rare, the minority variants that occurred in coding regions were non-synonymous, but with very low frequency alleles. The N-7 isolate harbored a 3.39% frequency of a non-synonymous variant in UL14 (Fig. 4), and another at 2.4% frequency in US6. The minority allele for UL14 encodes a valine to methionine change at residue 109 (V109M), which exists as the dominant allele in several independent isolates of HSV-1 (strains F (USA), CR38 (China), and R62 (South Korea))³¹. Most of the observed minority variants in each genome were present in intergenic regions, and adjoined tandem repeats or homopolymers (Supplementary Fig. 4). Length variations at tandem repeats and homopolymers are common between strains of HSV-1^{31,41,42}, and have been documented in select coding regions as well (e.g. UL30 polymerase and TK)⁴³.

Discussion

Here we provide the first-ever insights on viral inter- and intra-strain variation during familial transmission of HSV-1. These viruses appear indistinguishable in their phenotypes in culture and in an animal model of pathogenesis. Their clinical course in the infected father and son were similar, with both experiencing relatively severe gingivostomatitis at the time of initial infection (as noted by each mother), followed by six to eight recurrences per year thereafter. The randomly selected HSV-1 isolates from each individual were nearly identical to each other at the genomic level, despite being separated by over four decades, a horizontal transmission event, and multiple rounds of latency and reactivation. This exceptional level of genomic identity contrasts with that observed for *in utero* transmission of HCMV, which was found to include both a bottleneck and a subsequent expansion of viral diversity^{17,18,21,44}. However, HCMV and HSV-1 differ in their pathogenesis. HCMV is known to infect wider

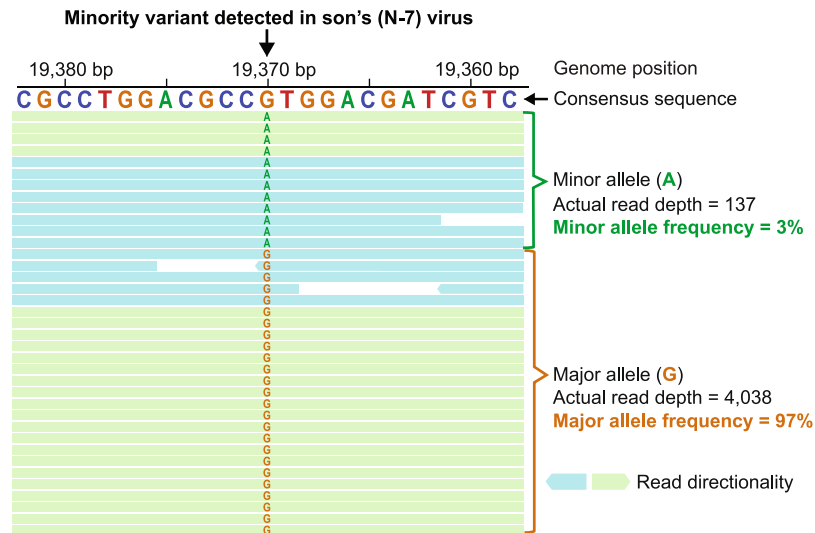


Figure 4. Intra-strain variation observed at a polymorphic locus in N-7 (son's) viral genome, in the gene UL14. A low frequency non-synonymous variation was observed in the N-7 (son's) viral genome at position 19,370, in the tegument protein UL14. This site has an A present in 3% of the viral sequence reads instead of the majority G allele (97%). The minority allele for UL14 encodes a valine to methionine change at residue 109; Met109 exists as the dominant allele in several independent isolates of HSV-1 (see text for details). While the UL14 coding sequence is encoded on the reverse strand of the reference genome for HSV-1, it is depicted here in forward orientation to enable codon reading from left to right. Actual read depth of each variant is indicated above. A subset of the alignment of Illumina sequencing reads to the N-7 consensus genome is shown here, with the position and consensus sequence shown in the top row. Sequence read orientation is depicted as aqua and green, with directional arrows. Areas with no letter shown have 100% agreement with the consensus nucleotide; the letters are left out for clarity.

range of cell types, host organs, and undergo widespread viral dissemination *in vivo* as compared to HSV-1⁴⁵. Furthermore, HCMV latency occurs in hematopoietic cells that undergo cell division, whereas the latent reservoir of HSV-1 in neurons does not undergo cell division. These distinctions may contribute to the observed differences in their level of intra-host variation.

The observation of nearly-perfect preservation of these viral genomes during familial transmission leaves open the question of where the observed diversity between clinical isolates of HSV-1 arises. One possibility (i) is that inter-strain diversity is the result of a slow mutation rate and variations that have accumulated over millions of years, with millions of clades of virus circulating around the planet. In this case the minority variants detected in the N-7 (son's) viral genome may represent the fodder for evolutionary selection. Similarities in HSV-1 isolates from the same geographic regions suggests that selection for fitness adaptations to the local human populations may have occurred. However these geographic patterns could also be the result of founder effects and sequestration of human movement in early millennia.

Another possible explanation (ii) for the observed inter-strain diversity of HSV-1 isolates is that horizontal transmission between unrelated hosts may induce greater selection for mutations than familial transmission. Sexual partners are unlikely to have matched immune selection for a newly transmitted HSV-1 isolate, and MHC alleles have been found to influence symptoms and severity of HSV disease^{46,47}. This hypothesis suggests that if a population of viruses are transmitted to a new unrelated host, then immune pressure over time in the new host could select for a viral population that differs from the one found in the original source.

Yet another possibility (iii) is that hosts with multiple HSV-1 infections or exposures may provide an opportunity for recombination to occur, with a concomitant rise in resulting variation of viruses shed by these individuals. Although rare, variations in the HSV-1 population(s) within an infected individual have been documented previously^{32,48,49}. Analyses of HSV-1 genomes by RFLP or full genome comparison suggests that recombination between viral genomes has been rampant over historical time^{31,50–54}. Finally, it is possible (iv) that different modes of transmission may affect the amount of variation transferred from a source individual, thus impacting the subsequent selection and fixation of these variants. For instance, it has been observed that shedding from lesion sites is associated with a high copy number of detectable viral genomes, whereas asymptomatic shedding has far fewer detectable genomes^{55,56}. This suggests that asymptomatic shedding, which is a recognized source of transmission between individuals^{57,58}, may represent a greater genetic bottleneck than transmission via contact with lesions.

In summary, using an animal model and HTS, we have shown that HSV-1 preserves its genetic and phenotypic characteristics during familial transmission. These approaches serve as a framework for future transmission studies using clinical isolates of HSV-1. Analysis of additional transmission events — particularly between unrelated individuals — will be necessary to assess how and where HSV-1 diversity arises.

Methods

Isolate acquisition and stock generation. Human Subjects were not recruited as part of this study. The samples described here were obtained through a virology laboratory at the Cincinnati Children's Hospital Medical Center. Informed consent was obtained from individuals contributing the viral samples. Once cultured, the viruses were coded and non-identifiable, and were not human tissues. HSV-1 isolates were collected from the father and the son on separate occasions, two years apart. For each isolate acquisition, a sterile swab was used to obtain virus from a lip vesicle during a recurrence. The swab was placed in sterile media and transported on ice to the laboratory, where a portion of the virus containing media was directly absorbed onto a rabbit skin cell (RSC) monolayer. This first passage was aliquoted and used to generate a "Pass 2" stock that was utilized for all phenotypic studies. These new isolates were named according to recent recommendations, as outlined by Kuhn *et al.* (23). We use shortened forms of these names, R-13 (father) and N-7 (son) throughout the manuscript. The full names for these isolates are HSV-1/Cincinnati, USA/1995/R-13 and HSV-1/Cincinnati, USA/1993/N-7.

Animal studies. All procedures in mice were performed as approved by the Children's Hospital Institutional Animal Care and Use Committee (protocol# IACUC2013-0162) and were in compliance with the *Guide for the Care and Use of Laboratory Animals*. Animals were housed in American Association for Laboratory Animal Care-approved quarters. Male outbred Swiss Webster mice (22–25 grams in weight) obtained from Harlan Laboratories (now Envigo) were used for these studies. Prior to infection, mice were anaesthetized by intraperitoneal injection of sodium pentobarbital (50 mg/kg of body weight). A 10 μ L drop containing 2×10^5 PFU was placed onto each scarified corneal surface as detailed previously²⁹.

Virus replication *in vivo*. At the indicated times post infection (pi) infected mice were euthanized and tissues, including eyes, trigeminal ganglia (TG), and brains from three mice from each inoculation group were individually assayed for virus as previously detailed²⁴. Isolation and quantification of total DNA from TG and quantification of total viral genomes by real time PCR using primers to the thymidine kinase region was performed as detailed previously⁵⁹.

Reactivation studies. Latent HSV was induced to reactivate in the ganglia of mice *in vivo* using hyperthermic stress (HS) (22). At 22 hours post induction, TG were assayed for infectious virus as detailed previously^{24,25}. For *in vitro* explant reactivation studies, latently infected ganglia were aseptically removed and placed into MEM supplemented with 5% newborn calf serum. These were incubated at 37 °C in a 5% CO₂ incubator. At the indicated times post explant, ganglia were homogenized and assayed for infectious virus as for reactivation *in vivo*⁶⁰.

Antibodies and Immunohistochemistry. Whole TG were fixed in 0.5% formaldehyde for 2 hours, rinsed in PBS and post fixed in DENT's fixative (80% methanol, 20% dimethylsulfoxide (DMSO)). Whole ganglia immunohistochemistry utilized a primary rabbit anti-HSV (AXL237, Accurate) at a 1:1000 dilution, and a secondary HRP-labeled goat anti-rabbit (Vector), at a 1:750 dilution. These methods have been detailed extensively in previous reports^{26,27,59,61,62}.

Virus culture and DNA isolation for HTS. Master stocks of the virus were prepared by infecting MRC-5 (ATCC[®], CCL-171) human fetal lung fibroblast cells grown in Eagle's Minimum Essential Media (EMEM:Sigma-Aldrich) at an MOI of 0.1. Viral stocks were harvested when significant cytopathic effect (CPE) was observed. Master stocks were tittered by limiting dilution assay on Vero cells (ATCC[®], CCL-171) to calculate plaque-forming units (PFU) for each stock. For isolation of viral DNA, MRC-5 cell cultures were infected at a multiplicity of infection (MOI) of 10 and viral genomic DNA (gDNA) was isolated using previously described methods for the isolation of viral nucleocapsid DNA⁶³.

Southern blot. The genomic structures of the clinical isolates were analyzed by DNA (Southern) blot analysis^{64,65} and compared to four commonly-used HSV-1 strains (17syn +, McKrae, KOS(M), and F) and 10 unrelated clinical isolates. Genomic DNA was cleaved with BamHI, transferred to 0.8% agarose gels, blotted, and probed with a cosmid clone insert that was ³²P-labeled using a Rediprime kit from Amersham. The cosmid clone insert matches the HSV-1 strain 17 genome from positions 24,698 to 64,405 (~40 kb). Blots were developed and analyzed on a Storm phosphorimager and quantified with GelQuantNet software.

Illumina high-throughput sequencing. Sequencing libraries for each of the isolates were prepared using the Illumina TruSeq Nano DNA Sample Prep Kit, according to the manufacturer's recommended protocol for sequencing of genomic DNA. The target DNA fragment size selected for library construction was 550 base pairs (bp). All the samples were sequenced on an in-house Illumina MiSeq using version 3 chemistry to obtain paired-end sequence fragments of 300 \times 300 bp. Base calling and image analysis was performed with the MiSeq Control Software (MCS) version 2.3.0.

De novo assembly of consensus genomes. HSV-1 genomes were assembled using a recently described viral genome assembly (VirGA) workflow³⁰. Briefly, VirGA combines quality control preprocessing of reads, *de novo* assembly, genome linearization and annotation, and post-assembly quality assessments. HSV-1 strain 17 (GenBank Accession: JN555585) was used as comparator for the reference-guided portion of viral genome assembly in VirGA. GenBank accessions are listed in Table 1.

Consensus genome comparison and phylogenetic analysis. Trimmed versions of the genomes lacking the terminal repeats were used for consensus genome comparison and for intra-strain polymorphism detection (below). We used the trimmed format of the genome to avoid over representation of the repeat regions

during comparison. Clustalw2 (43) was used to construct pairwise global nucleotide alignments between whole genome sequences, and pairwise global amino acid alignments between open reading frames. These alignments were utilized by downstream custom Python scripts to calculate percent identity, protein differences, and genomic DNA variation between samples. Phylogenetic networks were constructed using SplitsTree⁴⁶⁶ using the uncorrected_P distance, and all gaps were ignored. GenBank accession numbers and publications describing previously sequenced isolates are as follows: 17 (JN555585, NC_001806)^{12,13}; F (GU734771)^{36,67}; H129 (GU734772)^{36,68}; KOS (JQ673480, JQ780693)^{69,70}; McKrae (JQ730035, JX142173)^{37,71,72}; HF10 (DQ889502)⁷³; KOS63 (KT425110), KOS79 (KT425109)³²; India (KJ847330); L2 (KT780616), SC16⁷⁴; MacIntyre (KM222720)⁷⁵; CJ970 (JN420341.1), CJ311 (JN420338.1), 134 (JN4000093.1)⁷⁶; RE (KF498959); 160/1982 (LT594192), 132/1998 (LT594457), 1319/2005 (LT594108), 1394/2005 (LT594111), 66/2007 (LT594110), 20/2007 (LT594109), 369/2007 (LT594112), 2158/2007 (LT594106), 3083/2008 (LT594107), 172/2010 (LT594105)⁷⁷; CR38 (HM585508), EO3 (HM585509), E06 (HM585496), E07 (HM585497), E08 (HM585498), E10 (HM585499), E11 (HM585500), E12 (HM585501), E13 (HM585502), E14 (HM585510), E15 (HM585503), E19 (HM585511), E22 (HM585504), E23 (HM585505), E25 (HM585506), E35 (HM585507), R11 (HM585514), R62 (HM585515), S23 (HM585512), S25 (HM585513)³¹.

Intra-strain minority-variant detection. Minority variant loci or polymorphic positions within each consensus genome were detected using parameters that aimed to distinguish truly polymorphic sites from errors produced during sequencing or during *de novo* assembly of these genomes. VarScan v2.2.11⁷⁸ was used to detect variants present within each consensus genome. To aid in differentiating true variants from potential sequencing errors⁷⁹, the following variant calling parameters³⁹ were used: minimum variant allele frequency ≥ 0.02 ; base call quality ≥ 20 ; read depth at the position ≥ 100 ; independent reads supporting minor allele ≥ 5 . Polymorphisms with directional strand bias $\geq 90\%$ were excluded. The variants obtained from VarScan were then mapped back to the genome to understand their distribution and mutational impact using SnpEff and SnpSift^{80,81}. Polymorphic positions were visually assessed and hand-curated (Supplementary Table 1) to label those that bordered homopolymeric tracts or short sequence repeats (tandem repeats), since both can induce local mis-alignment of reads and lead to imprecise local polymorphism detection.

Data Availability. Viral genomes sequenced during this study are included in GenBank under the following Accessions: KY922718, R-13 (full name: HSV-1/Cincinnati, USA/1995/R-13; KY922719, N-7 (full name: HSV-1/Cincinnati, USA/1993/N-7). All other data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

References

- Wald, A. & Corey, L. Persistence in the population: epidemiology, transmission. in *Human Herpesviruses: Biology, Therapy, and Immunophylaxis* (eds Arvin, A. et al.) (Cambridge University Press, 2007).
- Corey, L. & Spear, P. G. Infections with Herpes Simplex Viruses. *N. Engl. J. Med.* **314**, 749–757 (1986).
- Whitley, R. J. Herpes simplex encephalitis: Adolescents and adults. *Antiviral Res.* **71**, 141–148 (2006).
- Park, P. J. et al. Corneal lymphangiogenesis in herpetic stromal keratitis. *Surv. Ophthalmol.* **60**, 60–71 (2015).
- Kimberlin, D. W. Herpes Simplex Virus Infections of the Newborn. *Semin. Perinatol.* **31**, 19–25 (2007).
- Corey, L. & Wald, A. Maternal and Neonatal HSV Infections. *N. Engl. J. Med.* **361**, 1376–1385 (2009).
- Corey, L. Synergistic copathogens—HIV-1 and HSV-2. *N. Engl. J. Med.* **356**, 854–856 (2007).
- Corey, L., Wald, A., Celum, C. L. & Quinn, T. C. The effects of herpes simplex virus-2 on HIV-1 acquisition and transmission: a review of two overlapping epidemics. *J. Acquir. Immune Defic. Syndr.* **1999** **35**, 435–445 (2004).
- Glynn, J. R., Biraro, S. & Weiss, H. A. Herpes simplex virus type 2: a key role in HIV incidence. *AIDS Lond. Engl.* **23**, 1595–1598 (2009).
- Hayward, G. S., Frenkel, N. & Roizman, B. Anatomy of herpes simplex virus DNA: strain differences and heterogeneity in the locations of restriction endonuclease cleavage sites. *Proc. Natl. Acad. Sci.* **72**, 1768–1772 (1975).
- Wiel, H. et al. Disseminated neonatal herpes simplex virus infection acquired from the father. *Eur. J. Pediatr.* **144**, 56–57 (1985).
- McGeoch, D. J. et al. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* **69**, 1531–74 (1988).
- McGeoch, D. J., Dolan, A., Donald, S. & Rixon, F. J. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. *J. Mol. Biol.* **181**, 1–13 (1985).
- Wang, G. P., Sherrill-Mix, S. A., Chang, K.-M., Quince, C. & Bushman, F. D. Hepatitis C Virus Transmission Bottlenecks Analyzed by Deep Sequencing. *J. Virol.* **84**, 6218–6228 (2010).
- Li, H. et al. Elucidation of Hepatitis C Virus Transmission and Early Diversification by Single Genome Sequencing. *PLOS Pathog.* **8**, e1002880 (2012).
- Masharsky, A. E. et al. A Substantial Transmission Bottleneck among Newly and Recently HIV-1-Infected Injection Drug Users in St Petersburg, Russia. *J. Infect. Dis.* **201**, 1697–1702 (2010).
- Renzette, N. et al. Rapid intrahost evolution of human cytomegalovirus is shaped by demography and positive selection. *PLoS Genet.* **9**, 1–14 (2013).
- Renzette, N., Gibson, L., Jensen, J. D. & Kowalik, T. F. Human cytomegalovirus intrahost evolution—a new avenue for understanding and controlling herpesvirus infections. *Curr. Opin. Virol.* **8**, 109–115 (2014).
- Andino, R. & Domingo, E. Viral quasispecies. *Virology* **479–480**, 46–51 (2015).
- Lauring, A. S. & Andino, R. Quasispecies Theory and the Behavior of RNA Viruses. *PLoS Pathog.* **6**, e1001005 (2010).
- Renzette, N., Bhattacharjee, B., Jensen, J. D., Gibson, L. & Kowalik, T. F. Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants. *PLoS Pathog.* **7**, 1–14 (2011).
- Sawtell, N. M. & Thompson, R. L. De Novo Herpes Simplex Virus VP16 Expression Gates a Dynamic Programmatic Transition and Sets the Latent/Lytic Balance during Acute Infection in Trigeminal Ganglia. *PLOS Pathog.* **12**, e1005877 (2016).
- Stevens, J. G. & Cook, M. L. Latent herpes simplex virus in spinal ganglia of mice. *Science* **173**, 843–845 (1971).
- Sawtell, N. M. & Thompson, R. L. Rapid *in vivo* reactivation of herpes simplex virus in latently infected murine ganglionic neurons after transient hyperthermia. *J. Virol.* **66**, 2150–2156 (1992).
- Sawtell, N. M. & Thompson, R. L. Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. *J. Virol.* **66**, 2157–2169 (1992).

26. Thompson, R. L. & Sawtell, N. M. The herpes simplex virus type 1 latency associated transcript locus is required for the maintenance of reactivation competent latent infections. *J. Neurovirol.* **17**, 552–558 (2011).
27. Sawtell, N. M. Quantitative Analysis of Herpes Simplex Virus Reactivation *In Vivo* Demonstrates that Reactivation in the Nervous System Is Not Inhibited at Early Times Postinoculation. *J. Virol.* **77**, 4127–4138 (2003).
28. Sawtell, N. M. The probability of *in vivo* reactivation of herpes simplex virus type 1 increases with the number of latently infected neurons in the ganglia. *J. Virol.* **72**, 6888–6892 (1998).
29. Sawtell, N. M. Detection and quantification of the rare latently infected cell undergoing herpes simplex virus transcriptional activation in the nervous system *in vivo*. *DNA Viruses Methods Protoc.* **292**, 57–72 (2005).
30. Parsons, L. R. *et al.* Rapid genome assembly and comparison decode intrastrain variation in human alphaherpesviruses. *mBio* **6**, (2015).
31. Szpara, M. L. *et al.* Evolution and diversity in human herpes simplex virus genomes. *J. Virol.* **88**, 1209–27 (2014).
32. Bowen, C. D. *et al.* Viral forensic genomics reveals the relatedness of classic herpes simplex virus strains KOS, KOS63, and KOS79. *Virology* **492**, 179–186 (2016).
33. Pandey, U. *et al.* DNA from dust: Comparative genomics of large DNA viruses in field surveillance samples. *mSphere* **1**, e00132–16 (2016).
34. Norberg, P. *et al.* A genome-wide comparative evolutionary analysis of herpes simplex virus type 1 and varicella zoster virus. *PLoS One* **6**, 1–8 (2011).
35. Grose, C. Pangaea and the Out-of-Africa Model of Varicella-Zoster Virus Evolution and Phylogeography. *J. Virol.* **86**, 9558–9565 (2012).
36. Szpara, M. L., Parsons, L. & Enquist, L. W. Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. *J. Virol.* **84**, 5303–13 (2010).
37. Watson, G. *et al.* Sequence and comparative analysis of the genome of HSV-1 strain McKrae. *Virology* **433**, 528–37 (2012).
38. Wagner, L. M. & DeLuca, N. A. Temporal Association of Herpes Simplex Virus ICP4 with Cellular Complexes Functioning at Multiple Steps in PolII Transcription. *PLOS ONE* **8**, e78242 (2013).
39. Depledge, D. P. *et al.* Deep Sequencing of Viral Genomes Provides Insight into the Evolution and Pathogenesis of Varicella Zoster Virus and Its Vaccine in Humans. *Mol. Biol. Evol.* **31**, 397–409 (2014).
40. Renzette, N., Pfeifer, S. P., Matuszewski, S., Kowalik, T. F. & Jensen, J. D. On the Analysis of Intrahost and Interhost Viral Populations: Human Cytomegalovirus as a Case Study of Pitfalls and Expectations. *J. Virol.* **91**, e01976–16 (2017).
41. Maertzdorf, J. *et al.* Amplification of reiterated sequences of herpes simplex virus type 1 (HSV-1) genome to discriminate between clinical HSV-1 isolates. *J. Clin. Microbiol.* **37**, 3518–3523 (1999).
42. Deback, C. *et al.* Utilization of microsatellite polymorphism for differentiating herpes simplex virus type 1 strains. *J. Clin. Microbiol.* **47**, 533–40 (2009).
43. Burrell, S., Deback, C., Agut, H. & Boutolleau, D. Genotypic Characterization of UL23 Thymidine Kinase and UL30 DNA Polymerase of Clinical Isolates of Herpes Simplex Virus: Natural Polymorphism and Mutations Associated with Resistance to Antivirals. *Antimicrob. Agents Chemother.* **54**, 4833–4842 (2010).
44. Renzette, N. *et al.* Limits and patterns of cytomegalovirus genomic diversity in humans. *Proc. Natl. Acad. Sci. USA*, <https://doi.org/10.1073/pnas.1501880112> (2015).
45. Muller, W., Jones, C. & Koelle, D. Immunobiology of Herpes Simplex Virus and Cytomegalovirus Infections of the Fetus and Newborn. *Curr. Immunol. Rev.* **6**, 38–55 (2010).
46. Orr, M. T. *et al.* Inhibition of MHC Class I Is a Virulence Factor in Herpes Simplex Virus Infection of Mice. *PLoS Pathog.* **1**, 62–71 (2005).
47. Lekstrom-Himes, J. A. *et al.* Association of Major Histocompatibility Complex Determinants with the Development of Symptomatic and Asymptomatic Genital Herpes Simplex Virus Type 2 Infections. *J. Infect. Dis.* **179**, 1077–1085 (1999).
48. Umene, K. & Yoshida, M. Reiterated sequences of herpes simplex virus type 1 (HSV-1) genome can serve as physical markers for the differentiation of HSV-1 strains. *Arch. Virol.* **106**, 281–299 (1989).
49. Heller, M., Dix, R. D., Baringer, J. R., Schachter, J. & Conte, J. E. Jr. Herpetic proctitis and meningitis: recovery of two strains of herpes simplex virus type 1 from cerebrospinal fluid. *J. Infect. Dis.* **146**, 584–588 (1982).
50. Bowden, R., Sakaoka, H., Donnelly, P. & Ward, R. High recombination rate in herpes simplex virus type 1 natural populations suggests significant co-infection. *Infect. Genet. Evol.* **4**, 115–23 (2004).
51. Sakaoka, H. *et al.* Genomic comparison of herpes simplex virus type 1 isolates from Japan, Sweden and Kenya. *J. Gen. Virol.* **68**, 749–764 (1987).
52. Sakaoka, H. *et al.* Quantitative analysis of genomic polymorphism of herpes simplex virus type 1 strains from six countries: studies of molecular evolution and molecular epidemiology of the virus. *J. Gen. Virol.* **75**, 513–27 (1994).
53. Norberg, P., Bergstrom, T., Rekabdar, E. & Lindh, M. Phylogenetic Analysis of Clinical Herpes Simplex Virus Type 1 Isolates Identified Three Genetic Groups and Recombinant Viruses. *J. Virol.* **78**, 10755–10764 (2004).
54. Kolb, A. W., Ané, C. & Brandt, C. R. Using HSV-1 genome phylogenetics to track past human migrations. *PLoS One* **8**, 1–9 (2013).
55. Tronstein, E. *et al.* Genital shedding of herpes simplex virus among symptomatic and asymptomatic persons with HSV-2 infection. *Jama* **305**, 1441–1449 (2011).
56. Johnston, C. *et al.* Virologic and Immunologic Evidence of Multifocal Genital Herpes Simplex Virus 2 Infection. *J. Virol.* **88**, 4921–4931 (2014).
57. Langenberg, A. G., Corey, L., Ashley, R. L., Leong, W. P. & Straus, S. E. A prospective study of new infections with herpes simplex virus type 1 and type 2. Chiron HSV Vaccine Study Group. *N. Engl. J. Med.* **341**, 1432–1438 (1999).
58. Corey, L. *et al.* Once-daily valacyclovir to reduce the risk of transmission of genital herpes. *N. Engl. J. Med.* **350**, 11–20 (2004).
59. Sawtell, N. M., Thompson, R. L. & Haas, R. L. Herpes Simplex Virus DNA Synthesis Is Not a Decisive Regulatory Event in the Initiation of Lytic Viral Protein Expression in Neurons *In Vivo* during Primary Infection or Reactivation from Latency. *J. Virol.* **80**, 38–50 (2006).
60. Sawtell, N. M. & Thompson, R. L. Comparison of Herpes Simplex Virus Reactivation in Ganglia *In Vivo* and in Explants Demonstrates Quantitative and Qualitative Differences. *J. Virol.* **78**, 7784–7794 (2004).
61. Thompson, R. L., Shieh, M. T. & Sawtell, N. M. Analysis of Herpes Simplex Virus ICP0 Promoter Function in Sensory Neurons during Acute Infection, Establishment of Latency, and Reactivation *In Vivo*. *J. Virol.* **77**, 12319–12330 (2003).
62. Thompson, R. L., Preston, C. M. & Sawtell, N. M. De novo synthesis of VP16 coordinates the exit from HSV latency *in vivo*. *PLoS Pathog.* **5**, 1–14 (2009).
63. Szpara, M. Isolation of herpes simplex virus nucleocapsid DNA. in *Herpes Simplex Virus* (eds Diefenbach, R. J. & Fraefel, C.) **1144**, 31–41 (Springer New York, 2014).
64. Thompson, R. L., Wagner, E. K. & Stevens, J. G. Physical location of a herpes simplex virus type-1 gene function(s) specifically associated with a 10 million-fold increase in HSV neurovirulence. *Virology* **131**, 180–192 (1983).
65. Javier, R. T., Izumi, K. M. & Stevens, J. G. Localization of a herpes simplex virus neurovirulence gene dissociated from high-titer virus replication in the brain. *J. Virol.* **62**, 1381–1387 (1988).
66. Huson, D. H. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**, 68–73 (1998).
67. Ejercito, P. M., Kieff, E. D. & Roizman, B. Characterization of herpes simplex virus strains differing in their effects on social behaviour of infected cells. *J. Gen. Virol.* **2**, 357–364 (1968).

68. Dix, R. D., McKendall, R. R. & Baringer, J. R. Comparative neurovirulence of herpes simplex virus type 1 strains after peripheral or intracerebral inoculation of BALB/c mice. *Infect. Immun.* **40**, 103–112 (1983).
69. Macdonald, S. J., Mostafa, H. H., Morrison, L. A. & Davido, D. J. Genome sequence of herpes simplex virus 1 strain KOS. *J. Virol.* **86**, 6371–6372 (2012).
70. Smith, K. O. Relationship Between the Envelope and the Infectivity of Herpes Simplex Virus. *Exp. Biol. Med.* **115**, 814–816 (1964).
71. Macdonald, S. J., Mostafa, H. H., Morrison, L. A. & Davido, D. J. Genome sequence of herpes simplex virus 1 strain McKrae. *J. Virol.* **86**, 9540–9541 (2012).
72. Williams, L. E., Nesburn, A. B. & Kaufman, H. E. Experimental induction of disciform keratitis. *Arch. Ophthalmol.* **73**, 112–114 (1965).
73. Ushijima, Y. *et al.* Determination and analysis of the DNA sequence of highly attenuated herpes simplex virus type 1 mutant HF10, a potential oncolytic virus. *Microbes Infect.* **9**, 142–149 (2007).
74. Rastrojo, A., López-Muñoz, A. D. & Alcamí, A. Genome Sequence of Herpes Simplex Virus 1 Strain SC16. *Genome Announc.* **5**, e01392–16 (2017).
75. Szpara, M. L. *et al.* Genome sequence of the anterograde-spread-defective herpes simplex virus 1 strain MacIntyre. *Genome Announc.* **2**, (2014).
76. Kolb, A. W., Lee, K., Larsen, I., Craven, M. & Brandt, C. R. 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 (2016).
77. Pfaff, F., Groth, M., Sauerbrei, A. & Zell, R. Genotyping of herpes simplex virus type 1 (HSV-1) by whole genome sequencing. *J. Gen. Virol.* <https://doi.org/10.1099/jgv.0.000589> (2016).
78. Koboldt, D. C. *et al.* VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
79. Nakamura, K. *et al.* Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Res.* **39**, 1–13 (2011).
80. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* **6**, 80–92 (2012).
81. Cingolani, P. *et al.* Using *Drosophila melanogaster* as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front. Genet.* **3** (2012).

Acknowledgements

We thank Yoly Tafuri, Jacob Shreve, Lynn Enquist, and Lance Parsons for contributions to the early phases of this work, as well as members of the Thompson, Sawtell and Szpara labs for helpful feedback and discussion. Genomics work was supported by the Virus Pathogens Resource (ViPR), a Bioinformatics Resource Center (BRC) funded by NIAID, and by the Huck Institutes for the Life Sciences at the Pennsylvania State University (MLS). This project was funded, in part, under a grant with the Pennsylvania Department of Health using Tobacco CURE funds (MLS). The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions. Additional support was provided by 1R01AI093614 grants to NS and RT.

Author Contributions

U.P. prepared viruses for sequencing and performed computational comparisons. D.R. assembled the full-length genomes. R.T. performed Southern blot analyses. N.S. performed all animal pathogenesis work. M.S. and N.S. conceptualized the work. M.S. and U.P. wrote the manuscript; all authors contributed to its completion. Accession codes: GenBank Accessions: R-13 = KY922718, N-7 = KY922719.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-13936-6>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017