



Personalized viral genomic investigation of herpes simplex virus 1 perinatal viremic transmission with dual fatality

Mackenzie M. Shipley,^{1,2} Daniel W. Renner,^{1,2} Utsav Pandey,^{1,2} Bradley Ford,³ David C. Bloom,⁴ Charles Grose,⁵ and Moriah L. Szpara^{1,2}

¹Department of Biochemistry and Molecular Biology, ²Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA; ³Department of Pathology, University of Iowa, Iowa City, Iowa 52242, USA; ⁴Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, Gainesville, Florida 32610, USA; ⁵Division of Infectious Disease/Virology, University of Iowa, Iowa City, Iowa 52242, USA

Abstract Here we present a personalized viral genomics approach to investigating a rare case of perinatal herpes simplex virus 1 (HSV-1) transmission that ended in death of both mother and neonate. We sought to determine whether the virus involved in this rare case had any unusual features that may have contributed to the dire patient outcome. A pregnant woman with negative HerpeSelect antibody test underwent cesarean section at 30 wk gestation and died the same day. The premature newborn died 5 d later. Both individuals were found postmortem to have positive blood HSV-1 PCR tests. Using oligonucleotide enrichment and deep sequencing, we determined that viral transmission from mother to infant was nearly perfect at the consensus genome level. At the virus population level, 77% of minor variants (MVs) in the mother's blood also appeared on the neonate's skin, of which more than half were disseminated into the neonate's blood. We also detected nonmaternal MVs that arose de novo in the neonate's viral populations. Of note, one de novo MV in the neonate's skin virus induced a nonsynonymous mutation in the UL6 protein, which is a component of the portal that allows DNA entry into new progeny capsids. This case suggests that perinatal viremic HSV-1 transmission includes the majority of genetic diversity from the maternal virus population and that new, nonsynonymous mutations can occur after relatively few rounds of replication. This report expands our understanding of viral transmission in humans and may lead to improved diagnostic strategies for neonatal HSV-1 acquisition.

Corresponding author:
moriah@psu.edu

© 2019 Shipley et al. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial License, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.

Ontology terms: herpes simplex encephalitis; preeclampsia; severe viral infections

Published by Cold Spring Harbor Laboratory Press

doi:10.1101/mcs.a004382

[Supplemental material is available for this article.]

INTRODUCTION

Human herpesvirus, also known as herpes simplex virus (HSV), is a large double-stranded DNA virus (*Simplexvirus*, Alphaherpesvirinae, Herpesviridae) that is prevalent all over the world, with >60% of people infected with HSV type 1 (HSV-1) (Looker et al. 2015a). HSV type 2 (HSV-2) is less prevalent, but still affects ~19% of the global population (Looker et al. 2015b). The viral life cycle of HSV includes entry and amplification of viral progeny at the epithelial mucosa, which typically results in the formation of oral or genital lesions. The virus is also able to enter innervating axonal nerve endings, whereupon it establishes a persistent infection in sensory or sympathetic neuronal nuclei of the peripheral nervous system.

In these peripheral neurons, the virus is able to evade the host immune system as it is silently maintained, undergoing no viral replication during this latent phase (Roizman et al. 2013).

In addition to the painful oral and genital lesions that HSV can cause, it can also result in the development of infectious blindness because of herpes keratitis (Streilein et al. 1997; Farooq and Shukla 2012). In rare cases, adults can develop encephalitis that results in death or extreme neurological disease (Whitley 2006). Another sequela that may develop because of HSV infection is viremia, or a systemic infection in which the virus is present in the bloodstream. HSV viremia in adults has been documented and is usually associated with severe primary orofacial or genital symptoms in immune-competent patients and occasionally with reactivations in immunocompromised patients (Johnston et al. 2008; Berrington et al. 2009). HSV viremia is more common during viral infection of neonates (Kimura et al. 1991).

Neonatal HSV is typically defined as the acquisition of viral infection within 28 d of birth (Corey and Wald 2009). The predicted burden of neonatal HSV disease in the Americas is ~1 case per 5000 live births (Looker et al. 2017). The majority of neonatal HSV infections are acquired during birth (Whitley 1986; Grose 1994; Kimberlin 2007; Corey and Wald 2009). Less common routes of neonate infection include viral transmission in utero and postnatal acquisition from exposure to HSV via family members or caregivers. Neonatal HSV infections can be classified into three categories based on the initial disease presentation: skin, eye, and mouth (SEM); disseminated disease; or central nervous system (CNS) disease (Whitley 1986; Kimberlin 2007; Corey and Wald 2009). Disseminated neonatal HSV infections have the highest rates of mortality (~85%), whereas CNS disease results in ~50% mortality, if left untreated (Corey and Wald 2009). Newborns with diagnosed HSV infection are typically treated with intravenous acyclovir for 3 wk at a dose of 60 mg/kg/d (Kimberlin et al. 2001a,b). There are several factors that may contribute to an increased probability of neonatal transmission in HSV-infected individuals. These include first-episode genital HSV disease in the mother, asymptomatic viral shedding at the time of delivery, and route of delivery (Brown et al. 2003). Cesarean section performed when a risk of vaginal transmission is inferred typically lowers the chances of neonatal HSV acquisition during delivery (Brown et al. 2003).

This report focuses on a specific case of perinatal HSV-1 transmission that was accompanied by dual mortality. The HSV etiology and subsequent transmission event were unanticipated because they occurred on the same day that the HerpeSelect HSV-1/HSV-2 IgG antibody titer of the pregnant female was returned with a negative result. These data provide the first-ever genomic insights into the scope of viral genetic diversity transmitted from mother to neonate during a viremic infection.

RESULTS

Patient History and Clinical Presentation

The adult woman was a 41-yr-old female in week 30 of gestation. She had been transferred from an outside hospital with a diagnosis of severe preeclampsia and underwent an emergency cesarean section. Because of signs of hepatitis, she also had screening laboratory tests for HSV-1, HSV-2, cytomegalovirus (CMV), varicella zoster virus (VZV), parvovirus, hepatitis A, hepatitis B, and hepatitis C. The results of an HSV-1/HSV-2 antibody test were returned within 4 h and showed no IgG antibody present (HerpeSelect, Focus Diagnostics). The patient died later on the day of admission with hepatic and kidney failure. An HSV-specific polymerase chain reaction (PCR) test was requested on the patient's blood and processed by an outside reference laboratory, producing a positive HSV-1 result 4 d later (Fig. 1A). There were no comorbidities detected; the screening laboratory tests were positive only for hepatitis B antibody, from prior exposure to the hepatitis B vaccine.

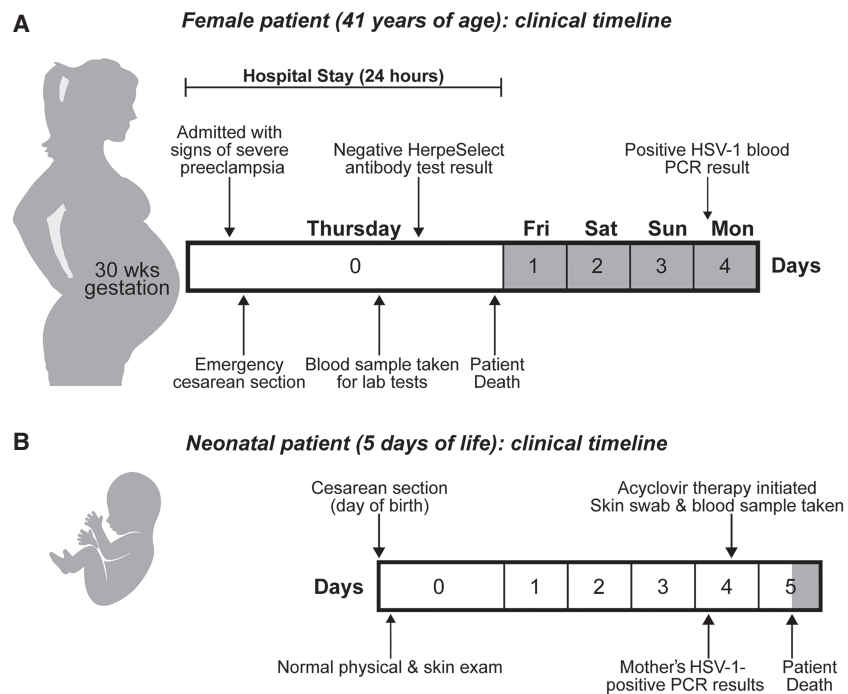


Figure 1. Timeline of clinical symptoms, HSV-1 detection, and dual fatality for mother and neonate. (A) The mother was admitted to the hospital with severe preeclampsia. An emergency cesarean section was performed, with the neonate at 30 wk gestation. A blood sample was collected for laboratory tests screening for HSV-1, HSV-2, cytomegalovirus (CMV), varicella zoster virus (VZV), parvovirus, hepatitis A, hepatitis B, and hepatitis C. The HerpeSelect IgG antibody test for HSV was negative. The mother died within 24 h. A PCR test on the mother's blood sample sent to a reference laboratory was returned with a positive HSV-1 result 4 d postmortem. (B) The neonate was delivered at 30 wk by emergency cesarean section and had a normal physical examination including normal skin appearance. When the mother's positive HSV-1 PCR result was returned on day 4, the neonate was immediately put on intravenous acyclovir therapy. At the time of acyclovir initiation, skin swab and blood samples were taken from the neonate for downstream testing. The neonate died the following day. The samples that were obtained from the neonate later tested positive for HSV-1 DNA by PCR. Note that boxes and arrows are not drawn to scale within each day but rather have been spaced out for legibility of text.

The newborn, although premature, had a standard physical inspection for week 30, including a normal skin examination—there were no visible lesions present on the neonate's skin. Following birth, the neonate was put on ampicillin/gentamicin treatment. Antibiotic treatment was stopped after 4 d, when bacterial cultures of the neonate's blood came back negative. When the mother's HSV-positive blood PCR test result was reported on day 4 of the infant's life, intravenous acyclovir therapy was initiated. A skin swab of the neonate's face and mouth and a blood sample collected at the time of acyclovir initiation tested positive in an HSV-specific PCR test. The infant died within a day of initiating acyclovir therapy (Fig. 1B). The neonate had severe liver failure at the time of death but never developed a skin rash. This dual mortality and availability of diagnostic specimens motivated a personalized viral genomics approach to investigate any potential viral features that may have contributed to this untimely outcome.

Near-Perfect Transmission of HSV-1 between Mother and Neonate at the Consensus Level

Viral genomic analysis was performed using a targeted genomic enrichment strategy that used oligonucleotide baits designed to capture viral DNA present in samples of the mother's

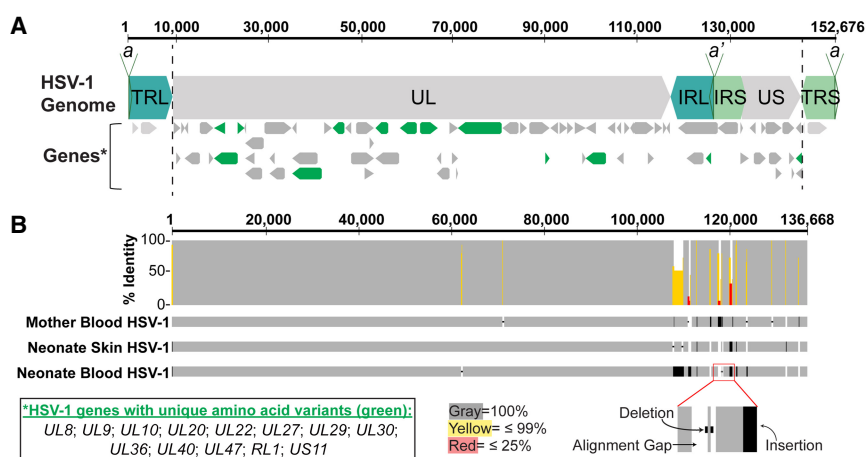


Figure 2. Nearly identical viral genomes exist in the mother and child at the consensus level. (A) Diagram of the HSV-1 genome and its genes. Overlapping genes are shown below the main diagram (arrows depict forward- vs. reverse-strand encoded genes), and highlighted genes (green) represent those that encode amino acid variants unique to these patient samples (see Table 2 for full list). Names of these genes are listed in an inset below B. Black dashed vertical lines indicate the terminal repeat regions that were excluded (or trimmed) for downstream genomic analyses (see Methods). (B) A histogram showing percent identity of a DNA alignment of viral genomes derived from the three clinical samples from mother and neonate. In the histogram, nucleotide identity is color-coded: 100% identity is gray, $\leq 99\%$ identity is yellow, and $\leq 25\%$ identity is red. To illustrate the genome-specific locations of these nonidentical sites, each genome is depicted as a horizontal gray bar (bottom), with gaps in the alignment of the three genomes (indels) shown as vertical or horizontal black bars. (UL) Unique long region, (US) unique short region, (TRL/IRL) terminal or internal repeat of the long region, (TRS/IRS) terminal or internal repeat of the short region, (a,a') cleavage/packaging signal. Identity graph was generated using Geneious.

blood, neonate's skin swab, and neonate's blood. Following next-generation sequencing (NGS) and viral genome assembly, we confirmed that both the mother and child were infected with HSV-1 (see Supplemental Table S1 for sequencing statistics). In comparing the viral populations present in the mother and the child, we observed that the HSV-1 genomes from each patient were nearly identical at the consensus level, with an average pairwise nucleotide identity of 98%, thus providing high confidence in the inferred route of transmission (Fig. 2). When we performed a pairwise comparison with each viral genome presented in this work, we found that the mother's HSV-1 and neonate's skin HSV-1 genomes were 98.9% identical, the mother's HSV-1 and neonate's blood HSV-1 genomes were 96.8% identical, and the neonate's skin and neonate's blood HSV-1 genomes were 97.4% identical at the consensus level. We identified 17 sites of genomic difference between the three clinical virus genomes at the consensus level. The majority of these variations were insertions or deletions (indels) that occurred in either noncoding regions of genes (U_51 and U_55), in the large structural repeat regions of the genome, or in intergenic regions. There were no single-nucleotide differences or variants (SNVs) detected between the three viral consensus genomes presented here. The remaining disagreements were uncalled bases or "Ns" present in repetitive areas of the genome. These data suggested that at the consensus level, the HSV-1 genomes from mother and neonate were nearly identical.

Transmission of Minor Variants in the Viral Population Is Highly Conserved between Mother and Child

Using the extremely deep sequencing coverage of these genomes (3000 \times –7500 \times ; see Supplemental Table S1), we next examined the genetic diversity within each viral population

that was sampled. To do this we examined the sequence read data for any alternative alleles in each viral population. Minor variants (MVs) are alternative alleles that exist at low frequencies (between 2% and 50% of the sequence reads at a given locus, or a frequency of ≥ 0.02 and < 0.50) below the consensus level in the viral population. We sought to detect whether any MVs were present, and if any of these alternative alleles were conserved following viral transmission (Fig. 3). We identified 39 MVs present in the viral population of the mother's blood. Eight of these MVs occurred in genes and caused nonsynonymous, synonymous, or frameshift mutations (insertion or deletion variants) (see Supplemental Table S2 for a full list). The remaining 31 MVs in the viral population of the mother's blood were all intergenic and mostly occurred in the repetitive regions of the genome (Fig. 3A). The viral population derived from the neonate's skin swab had a total of 42 MVs, of which 13 occurred in genes and 29 were intergenic (Supplemental Table S2). The viral population derived from the neonate's blood had the fewest number of MVs (28 total), of which eight occurred in genic regions of the genome.

To investigate whether perinatal transmission includes a bottleneck, we compared the nucleotide positions and frequencies of each MV found in the three viral populations using the mother's blood HSV-1 genome as an initial reference point (i.e., the donor population). The majority of the viral genetic diversity (77% of MVs) detected in the mother's viral population was transmitted to the neonate's skin, with the frequency of each MV being highly correlated in both mother and neonate (Fig. 3B). Slightly less of the mother's viral population diversity (59% of MVs) was found in the neonate's blood, although the MV frequencies were still highly correlated (Fig. 3B). Based on one possible route of infection from maternal blood to neonate's skin, and then into the neonate's blood, we also compared the potential transmission of viral genetic diversity from neonate's skin to blood. In this comparison, we found that 54.8% of MVs present in the neonate's skin viral population appeared in the neonate's blood, again with highly correlated frequencies (Fig. 3B). However in each comparison, we found several MVs that appeared not to be transmitted between the two hosts, as well as others that appeared to have arisen *de novo* (Table 1; Fig. 3B,C). Specifically, we observed 23 total MVs in both viral populations from the neonate that were not present in the mother's viral genome—the hypothesized donor population—at frequencies $\geq 2\%$ (0.02). These included a nonsynonymous G486W MV in the *UL6* gene at a frequency of 9% (0.09) (Table 1; Fig. 4), synonymous MVs present in *UL29* and *UL36*, many intergenic variants, and multiple disruptive frameshift mutations in *UL28*, *UL46*, *UL50*, *UL54*, and *UL55*. However, when we lowered the threshold ($\geq 0.2\%$ or 0.002) for MV detection to search for the presence of any of these apparently *de novo* MVs in the mother's blood HSV-1 population, we found that 14 of the 23 potentially *de novo* MVs actually existed in the mother's viral population at very low frequencies ($\geq 0.2\%$ but $< 2\%$) (Supplemental Table S2). Although distinguishing between biological MVs and sequencing artifacts is difficult at $< 1\%$, the existence of these MVs in the neonate's viral populations at increased frequencies suggested that they were transmitted from the mother's viral population rather than appearing *de novo*. Only $\sim 10\%$ of both viral populations in the neonate were made up of MVs that appeared to be *de novo*, including the nonsynonymous *UL6* MV, which is not detectable in the mother's viral population even at our lowest threshold of detection. Analysis of the viral populations present in the mother and neonate suggested that at the level of MVs, viral diversity was highly conserved between hosts, as well as between body compartments of the neonate (skin and blood).

Several of the MVs described above that were detected in all three viral populations have the potential to impact protein-coding functions of HSV-1. Specifically, all three virus populations from the mother and neonate contained genic MVs with $\geq 2\%$ (0.02) frequency in *UL30* and *US7* (Table 1). These shared MVs were indel variants that induce a frameshift in a protein-coding region. The indel MV that appeared most intriguing based on known protein

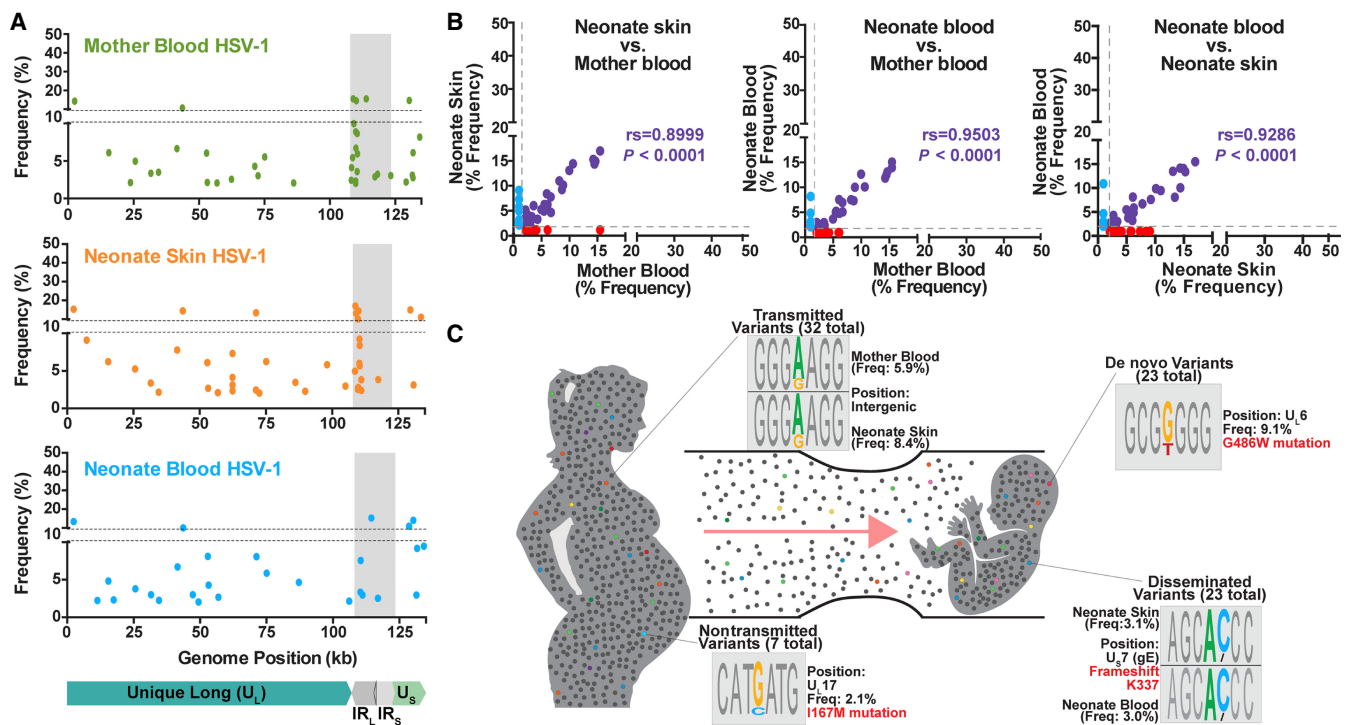


Figure 3. HSV-1 minor variants (MVs) are highly conserved between mother and neonate following perinatal transmission. (A) MVs identified in each of the three viral populations are plotted based on their location within the HSV-1 genome and frequency at each position. The average sequencing coverage depth of these MV loci is 2708 \times . Shaded areas on each graph indicate the large structural repeat regions (IR_L and IR_S)—see Figure 2A for greater detail. A small schematic of the trimmed HSV-1 genome is shown below A for reference. See Supplemental Table S2 for the classification of each MV shown here as nonsynonymous (genic), synonymous (genic), or intergenic. (B) Scatterplots depict the relative frequency of each matched MV detected in the mother’s blood, the neonate’s skin, and the neonate’s blood (see Methods for details). Matching MVs in two viral populations are compared, with the frequency of each MV in the “donor” population shown on the x-axis and the frequency of each MV in the recipient population shown on the y-axis. Points in purple indicate “transmitted” MVs that were present in both the donor and the recipient viral genome populations at frequencies $\geq 2\%$ (0.02). Points in red indicate potentially “nontransmitted” MVs that were present in the donor at $\geq 2\%$ (0.02), but were below this threshold in the recipient viral genome populations. Points in blue indicate potentially “de novo” MVs in the recipient viral genome populations; these were not detectable at the 2% (0.02) threshold in the donor viral genome populations. Spearman correlation analysis on the “transmitted” MVs (purple points) was performed in GraphPad Prism to obtain a correlation coefficient (rs) and corresponding significance value (P), shown in the center of each graph. Gray dashed lines show the MV detection threshold of 2% (0.02). (C) Model of perinatal transmission between mother and neonate, with highlighted examples of each class of MV identified in the three viral populations sampled here (Table 1). Transmitted variants (32 total) refer to MVs that were detected in both the mother’s blood HSV-1 sample and the neonate’s skin and/or blood HSV-1 samples (at frequencies $\geq 2\%$ or 0.02). Nontransmitted variants (seven total) refer to MVs that were detected only in the mother’s blood HSV-1 sample (at frequencies $\geq 2\%$ or 0.02). De novo variants (23 total) refer to MVs that were detected at frequencies $\geq 2\%$ (0.02) only in one or both of the neonate’s viral populations. The example shown in the U_L6 gene would result in a nonsynonymous mutation of glycine to tryptophan in the neonate’s skin HSV-1 (see Fig. 4 for more details). Disseminated variants (23 total) refer to MVs that appear to have spread from the neonate’s skin into the blood (at frequencies $\geq 2\%$ or 0.02). The example shown is a deletion variant in the U_S7 gene that codes for glycoprotein I, which results in a frameshift at AA position K337. Sequence motifs illustrate the penetrance of major and minor alleles at the highlighted (colored) nucleotide position, with surrounding nucleotides shown in gray. The nucleotide position and frequency of each MV example are given. Note that frequency values in this figure are depicted as percentages, in which the minimum variant allele frequency is $\geq 2\%$ (0.02). Data for this figure are a filtered set of high-confidence MVs that are a subset of all those detected—see Methods for details and Supplemental Table S2 for a full list of MVs.

Table 1. Details of minor variants (MVs) in the HSV-1 genomes obtained from mother and neonate that are shown in Figures 3C, 4, and 5

Isolate	Genome position	Major allele	Minor allele ^a	Minor allele frequency	Type of variation	Gene	Amino acid	MV conservation
Neonate's skin	7,373	G	T	9.09%	Nonsynonymous	<i>U_L6</i>	G486W	De novo
Mother's blood	23,780	G	C	2.13%	Nonsynonymous	<i>U_L17</i>	I167M	Nontransmitted
Mother's blood	56,884	G	GC	2.08%	Frameshift	<i>U_L30</i>	A1109	Transmitted (to neonate's skin and blood)
Neonate's skin	56,884	G	GC	2.10%	Frameshift	<i>U_L30</i>	A1109	Transmitted
Neonate's blood	56,884	G	GC	2.68%	Frameshift	<i>U_L30</i>	A1109	Transmitted
Mother's blood	110,565	A	G	5.96%	Intergenic	-	-	Transmitted (to neonate's skin only)
Neonate's skin	110,446	A	G	8.40%	Intergenic	-	-	Transmitted
Mother's blood	131,527	AC	A	3.12%	Frameshift	<i>U_S7</i>	K337	Transmitted (to neonate's skin and blood)
Neonate's skin	130,736	AC	A	3.13%	Frameshift	<i>U_S7</i>	K337	Transmitted
Neonate's blood	131,394	AC	A	2.96%	Frameshift	<i>U_S7</i>	K337	Transmitted

^aAdditional details for these and all other MVs detected in the HSV-1 genomes obtained from mother and neonate are included in Supplemental Table S2.

functions was a frameshift in *U_L30*, which codes for the viral polymerase subunit. This single-nucleotide insertion MV occurred in a C₈ homopolymer at A1109 (nucleotide position 56,884 in mother's blood-derived HSV-1 genome) (Table 1; Fig. 5). Another notable MV occurred in *U_S7*, which codes for glycoprotein I, one of the major glycoproteins involved in cell-to-cell spread of the virus and immune evasion (Crute and Lehman 1989; Dingwell et al. 1994, 1995; Bell et al. 1990). This *U_S7* variant is a frameshift caused by a deletion at K337 (nucleotide position 131,527 in mother's blood-derived HSV-1 genome) (Table 1; Fig. 3C). The untimely loss of both infected individuals precludes our ability to detect whether or how these variants may have impacted virus functions in the viral population in vivo.

Mother's and Neonate's HSV-1 Genomes Had North American/European Phylogeny with Distinct Mutations at the Protein-Coding Level

To determine how the three viral genomes involved in this case compared to other previously described strains of HSV-1, we performed a consensus level multiple sequence AA alignment with 48 other genomes representing the known global diversity of HSV-1 (see Supplemental Table S3 for published HSV-1 genomes, GenBank accession numbers, geographic origins of each strain, and related references). From this alignment, we identified 19 unique amino acid (AA) variants present in 13 viral proteins in the mother's and child's HSV-1 genomes. These AA variations have not been previously identified in any other known HSV-1 strains (Table 2). Based on BLOcks SUBstitution Matrix (BLOSUM) calculations, several of these unique AA variant substitutions have a low probability of occurrence, making their existence more intriguing. The BLOSUM score is based on odds ratios of all possible AA substitutions, calculated using conserved regions of sequence alignments from multiple protein families (Henikoff and Henikoff 1992). Examples of these unique AA variants include those in glycoprotein H (*U_L22*; R229I), which is required for viral fusion and entry into the host cell (Atanasiu et al. 2010), the single-stranded DNA binding protein called infected-cell protein 8 or ICP8 (*U_L29*; F214S), and the viral DNA polymerase (*U_L30*; L279R) (see Fig. 5).

We also examined the genetic relatedness of these three viral genomes to the other 48 HSV-1 genomes included in Supplemental Table S3 by creating a SplitsTree phylogenetic network using a whole-genome nucleotide alignment. We found that the mother's and

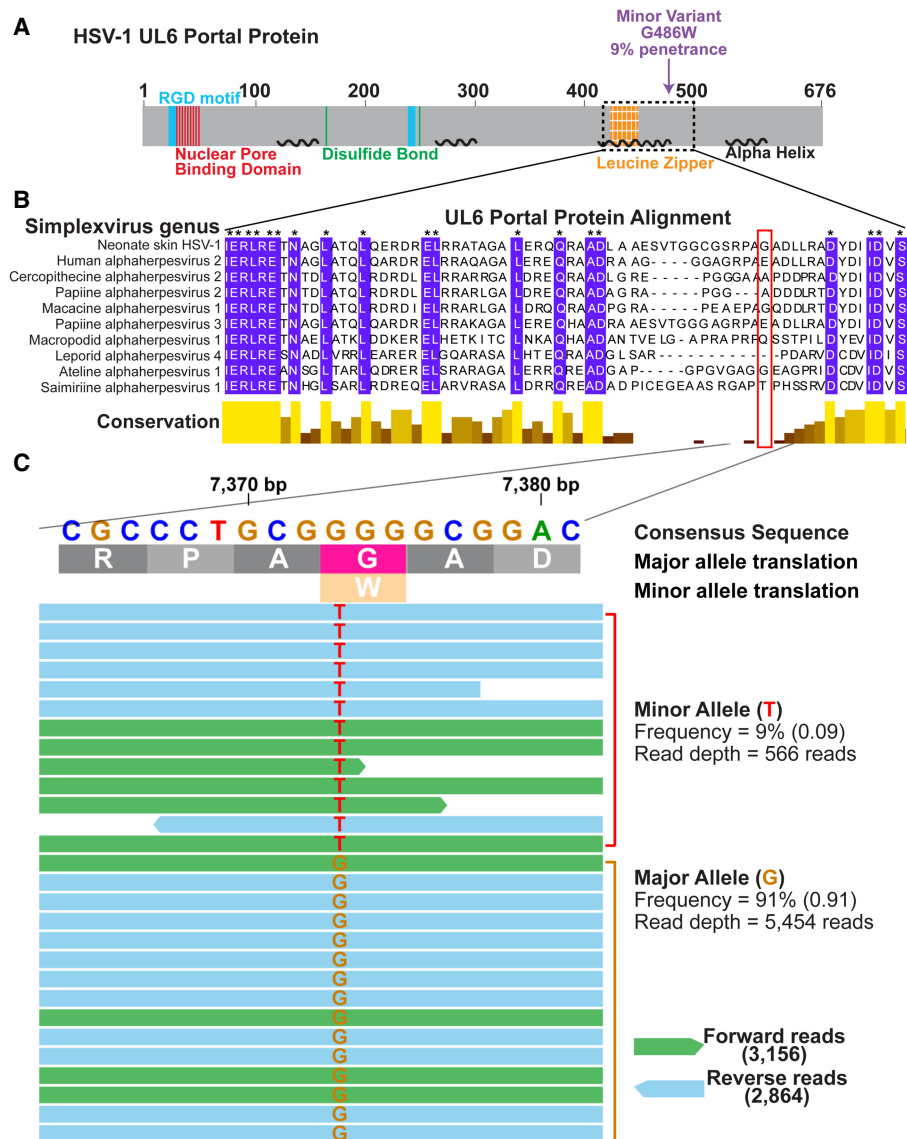


Figure 4. De novo minor variant (MV) in the UL6 portal protein of the neonate's skin HSV-1 genome. (A) Diagram of the UL6 portal protein of HSV-1 (676 amino acids). This graphic details the location of the UL₆ MV detected in the neonate's skin viral population at amino acid position 486, which results in a nonsynonymous mutation of glycine (G) to tryptophan (W). Predicted domains of the UL6 portal protein based on prior works are also highlighted (Newcomb et al. 2001; Nellissey et al. 2007; Liu et al. 2019). (B) The UL6 portal protein displays variable amino acid conservation across 10 virus species from the Simplexvirus genus, with an average pairwise identity of 62%. Specifically, amino acid position 486 (red box) is divergent across the Simplexvirus species. Amino acid positions that are 100% conserved across the 10 species are highlighted in purple and are denoted by a black asterisk above each position in the alignment. The histogram below the alignment plots the variable conservation across the protein alignment at each amino acid position. The amino acid positions shown for each species are listed to the right of the alignment. (C) The UL₆ MV encoding the G486W variant is highlighted at the nucleotide level of the neonate's skin viral population. A small subset of sequence reads that map to nucleotide position 7373 are shown, with the consensus level nucleotide sequence and its encoded amino acids at the top. The minor allele thymidine (T) at this position is present at a frequency of 9% (0.09), whereas the major allele guanine (G) is present at a frequency of 91% (0.91) (Table 1). Both forward (green) and reverse (blue) reads are shown to illustrate the bidirectional read support observed for both alleles.

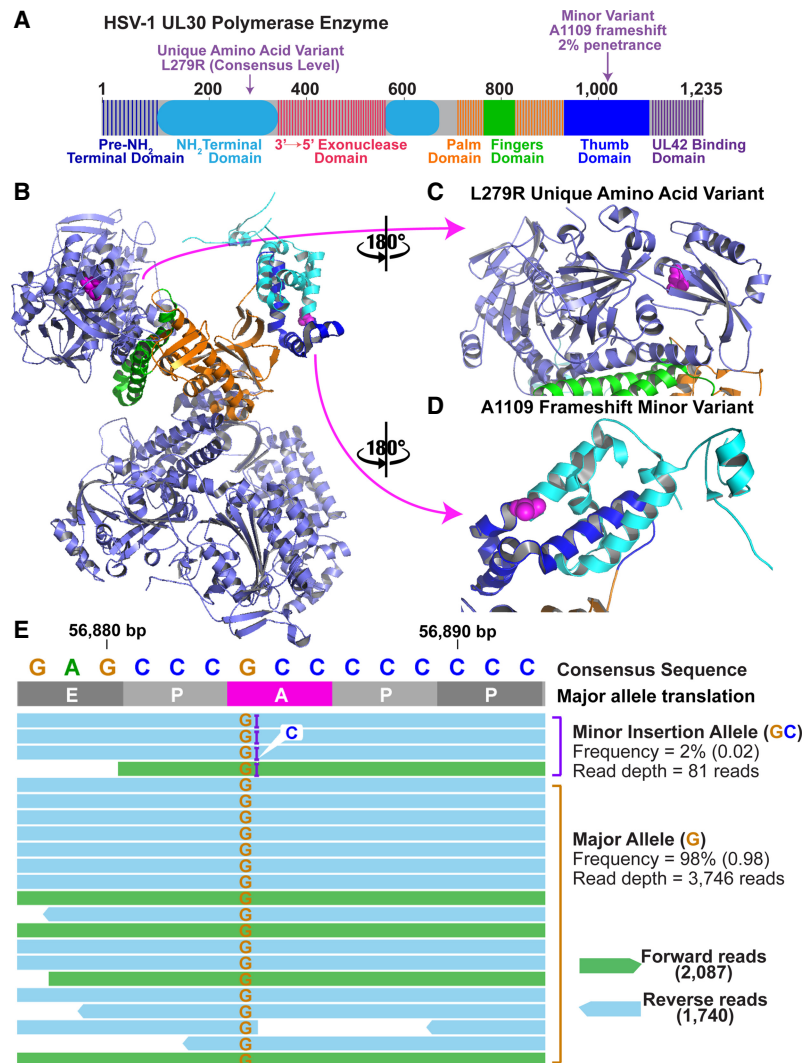


Figure 5. Unique variants in the UL30 viral polymerase enzyme are shared between all three clinical HSV-1 genomes from mother and neonate. (A) Protein diagram of the viral polymerase enzyme (1235 amino acids) encoded by the *UL30* gene of HSV-1, with highlighted domains of interest based on prior work (Andrei et al. 2000). This schematic details the location of the amino acid (AA) variant L279R that is found at the consensus level of all three viral genomes from mother and neonate and has never been observed before in any previously published HSV-1 isolates. The protein diagram also includes an MV insertion allele at A1109, which causes a frameshift mutation in the viral polymerase at a frequency of 2% in all three viral populations (Table 1). (B) Crystal structure ribbon diagram of the viral polymerase from HSV-1 strain KOS (Liu et al. 2006), with highlighted domains as follows: palm domain (AAs 701–766 and 826–956) in orange, fingers (AAs 767–825) in green, and thumb (AAs 957–1197) in blue. Colors match those shown on the carboxyl terminus of the UL30 diagram in A. Two sites of interest to the mother’s and neonate’s HSV-1 samples presented in this work are shown as magenta spheres. (C) 180° rotation across the vertical plane and zoom-in of the L279R AA variant (magenta) that is unique to the mother’s and neonate’s HSV-1 genomes. This variant occurs in the amino-terminal domain of the viral polymerase and is also highlighted in Tables 1 and 2. (D) 180° rotation across the vertical plane and zoom-in of the A1109 position (magenta), in which a subset of the sequencing reads encode a frameshift because of the presence of a single-nucleotide insertion. This MV occurs in the thumb domain of the viral polymerase. Parts of the structure that are colored cyan in B and D indicate amino acids that are downstream from A1109 and are thus most likely to be impacted by the frameshift variant. (E) The *UL30* MV encoding a frameshift insertion in all three viral populations is highlighted at the nucleotide level. Data shown here are from the mother’s blood HSV-1 genome population. A small subset of sequence reads that map to nucleotide position 56,884 are shown, with the consensus level nucleotide sequence and its encoded AAs at the top. The insertion of a “C” nucleotide occurs at a C₈ homopolymer and is present at a frequency of 2% (0.02), whereas the major allele guanine (G) is present at a frequency of 98% (0.98) (Table 1). Both forward (green) and reverse (blue) reads are shown to illustrate the bidirectional read support observed for both alleles.

Table 2. Unique amino acid (AA) variants found in all three clinical genomes from the mother and neonate that are not present in 48 other published strains of HSV-1

Gene	Gene product	AA alignment position	Wild-type allele ^a	Mother and neonate allele	BLOSUM80 score ^b
<i>U_L8</i>	DNA helicase/primase	41	N	T	0
		480	M	I	1
<i>U_L9</i>	Origin-binding protein, DNA replication	270	P	Q	-2
<i>U_L10</i>	Glycoprotein M (gM)	232	D	N	1
		456	G	E	-3
<i>U_L20</i>	Membrane protein	18	A	T	0
<i>U_L22</i>	Glycoprotein H (gH)	229	R	I	-3
<i>U_L27</i>	Glycoprotein B (gB)	79	K/P/T	Q	1/-2/-1
<i>U_L29</i>	ssDNA binding protein ICP8	214	F	S	-3
<i>U_L30</i>	DNA polymerase	279	L	R	-3
<i>U_L36</i>	VP1/2, tegument protein	491	L	I	1
<i>U_L40</i>	Ribonucleotide reductase, small subunit	114	L	I	1
		114	R	S	-1
<i>U_L47</i>	VP13/14—binds RNA	95	D	N	1
		100	Q	H	1
		114	R	S	-1
		278	T	P	-2
<i>U_L47</i>	VP13/14—binds RNA	279	S	N	0
		279	S	N	0
<i>R_L1</i>	ICP34.5, neurovirulence factor	84	D	N	1
<i>U_S11</i>	Tegument-associated phosphoprotein, posttranscriptional regulation	18	Y	S	-2

^aWild-type allele is found in all 48 other HSV-1 strains at each AA position.

^bBLOSUM80 identity score is the likelihood of AA substitution between protein sequences that have >80% identity (Altschul 1991; Henikoff and Henikoff 1992). More negative scores are less likely/frequently observed.

neonate's HSV-1 genomes clustered with other viral genomes in the North American/European geographical clade, consistent with their isolation from individuals in the United States (Fig. 6). Specifically, these clinical samples clustered on their own branch between HSV-1 McKrae (origin: Gainesville, Florida) (Williams et al. 1965) and HSV-1 strain 17 (origin: Glasgow, Scotland) (McGeoch et al. 1986, 1988). The deep-branch phylogeny of these three viral genomes suggested that they were either historically separated strains that had not been subject to any recent mixing or recombination, or that they had a unique geographic origin compared to the rest of the North American/European HSV-1 isolates that have been sampled and published thus far. A larger nucleotide alignment involving all available HSV-1 strains in GenBank (both partial and complete) produced similar results (data not shown).

DISCUSSION

After a thorough review of the literature, we were unable to find another report of HSV-1 transmission that occurred when the transmitting host was negative for HSV-1 IgG antibody by the sensitive HerpeSelect test but positive for HSV-1 DNA in the blood (viremia) on the same day. The goal of this study was to use both deep sequencing technology and comparative genomics to thoroughly investigate whether viral genomic features may have contributed to the severity of disease and the untimely demise of these patients, as has been done previously with other clinical HSV and varicella zoster virus cases (Santos et al. 1998; Wirtgart et al. 2006; Depledge et al. 2016). Here, we present a report detailing the transmission of

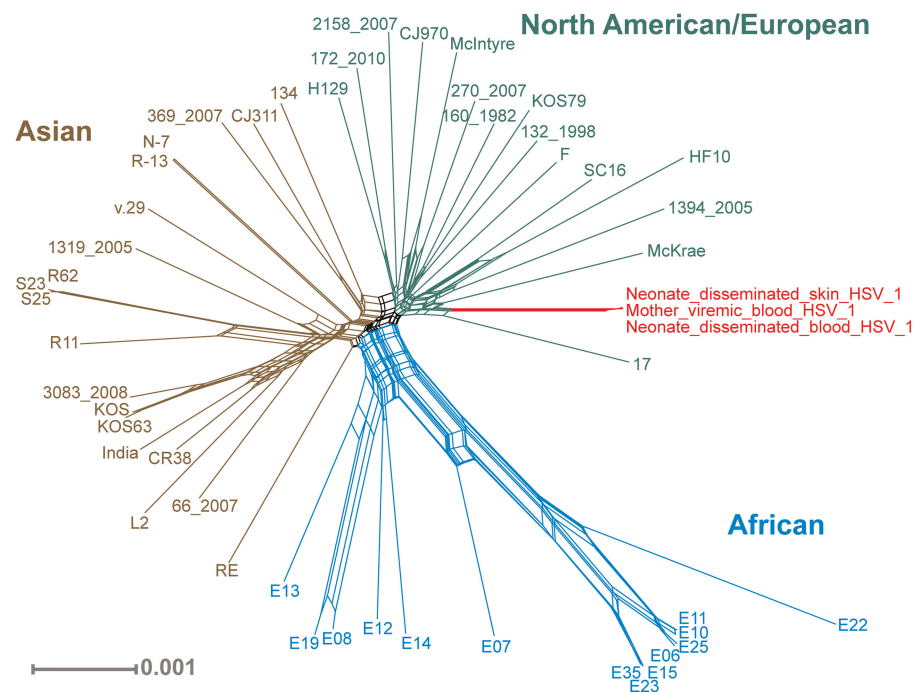


Figure 6. Network graph demonstrating the North American phylogeny of the mother’s and neonate’s HSV-1 genomes. The three clinical samples presented in this work are highlighted in red and fall within the North American/European phylogenetic clade (green). This graph-based network was generated by aligning the consensus sequences of the three viral genomes from mother and neonate with 48 published viral genomes that encompass the known global genetic diversity of HSV-1 (see Methods section and Supplemental Table S3 for names, accession numbers, and geographic origins of each strain). The Φ statistical test for recombination in SplitsTree4 reported $P = 0.0$, suggesting that recombination contributes to the phylogenetic network. Other geographical clades include Asian (brown) and African (blue). The scale bar represents 0.1% nucleotide divergence. N-7 and R-13 (upper left) are a pair of genomes from a previously reported case of familial transmission between a father and toddler son (Pandey et al. 2017).

HSV-1 from a mother with viremia to her newborn child, which ended in a dual fatality (Fig. 1). The pregnant female interacted frequently with young children, potentially enabling a late-stage exposure to oral HSV-1 infection from a child. It is highly likely that the woman’s primary HSV-1 infection acquired so late in pregnancy exacerbated the probability of neonatal transmission and severe infection (Brown et al. 2003). The mother’s primary HSV-1 infection most likely occurred within the 2 weeks prior to her emergency cesarean section, because of the lack of HSV-1 IgG antibody detected. The maternal-to-neonate transmission event is presumed to have occurred at the time of the cesarean section. However, we cannot rule out a possible in utero transmission event during the few days before cesarean section. Although estimating the time of the transmission event using the inferred mutation rate of HSV-1 was a possibility, there are multiple rates that have been reported in the literature, ranging from 1×10^{-8} mutations per base per infectious cycle to 1×10^{-4} mutations per base per year (Hall and Almy 1982; Drake and Hwang 2005; Firth et al. 2010). With these inconsistencies in mind, we did not endeavor to estimate the time of transmission in this case.

Through NGS analysis of uncultured viral genomes obtained from the mother’s blood, neonate’s skin, and neonate’s blood, we determined that HSV-1 transmission was nearly perfect at the level of the viral consensus genome (Fig. 2). The three viral consensus genomes were highly identical, with a few variations present in the large structural repeat regions of the genome or in noncoding regions of genes (i.e., U_S1 and U_S5). This level of genome

conservation might have been facilitated by the shared genetic background of the mother and child, as suggested by a recent study of the father-to-son transmission of HSV-1 (Pandey et al. 2017). Furthermore, the potential exposure to a large pool of virus, enabled by the mother's viremia and neonatal exposure to maternal blood during the emergency cesarean section, could have facilitated transmission and conservation of the majority of the viral population.

Although highly alike to one another at the consensus level, these HSV-1 genomes obtained from mother and child were found to have numerous unique AA variants present in key viral proteins (Table 2). Several of these mutations were considered to be rare or unlikely to occur based on BLOSUM calculations (Altschul 1991; Henikoff and Henikoff 1992). These included variants in glycoproteins M (gM, UL10) and H (gH, UL22), as well as in the single-stranded DNA binding protein ICP8 (UL29) and the viral polymerase subunit UL30 (Table 2; see Fig. 5 for UL30 variants). Glycoprotein H (gH) forms a heterodimer with glycoprotein L (gL), which in addition to being required for fusion and viral entry into the host cell (Hutchinson et al. 1992; Atanasiu et al. 2010) is also a target of anti-HSV antibodies (Peng et al. 1998). Although gH AA position 229 has not yet been reported as a site of viral escape from antibody recognition (Gompels 1991), the loss of positive charge due to the R229I mutation present in the mother's and child's HSV-1 genomes may well influence gH protein conformation, antibody epitope specificity, and subsequent viral escape (Table 2). The *UL29* gene encodes infected-cell protein 8 (ICP8), a single-stranded DNA binding protein that, together with host proteins, forms replication compartments and is imperative for viral DNA synthesis (Weller and Coen 2012). Although ICP8 AA position 214—which lies in the shoulder domain of the protein between β 11 and α 5—is not predicted to be critical for cooperative DNA binding based on the crystal structure (Mapelli et al. 2005), sites 206–214 are a predicted MHC supertype A24 epitope (FSRVIAEPF) (Pickett et al. 2012). This suggests that the F214S mutation present in the mother's and child's HSV-1 genomes may contribute to immune evasion through ablation of MHC recognition of this predicted immune epitope (Table 2). Despite the many interesting viral variants identified, we present only speculations regarding how these mutations may have influenced the overall pathogenicity of the virus in these individuals. To achieve predictive power of the effects that individual mutations have on virulence and disease outcomes, we must strive to obtain both phenotypic and genotypic data from HSV isolates, as well as improve our characterization of understudied viral proteins (Kolb et al. 2016; Akhtar et al. 2019).

Prior studies have demonstrated how MVs can have an effect on the fitness and evolution of multiple species of viruses including human immunodeficiency virus (HIV), influenza, human cytomegalovirus (HCMV), and varicella zoster virus (VZV) (Santos et al. 2000; Grose et al. 2004; Fischer et al. 2010; Renzette et al. 2013; Depledge et al. 2014; Bordería et al. 2015; Leonard et al. 2017). With the unusual case presented here, we wanted to determine whether a bottleneck in the viral population occurred during the perinatal transmission event. We used a conservative cutoff of $\geq 2\%$ (0.02 frequency) to detect with high confidence any MVs in the three viral genome populations sampled in the mother and neonate (Renzette et al. 2013; Depledge et al. 2014). When we examined the MVs that were present in each of the viral populations presented here, we found that most of the MVs from the mother's blood HSV-1 genome were transmitted to the neonate. We found that a similar group of MVs were conserved between the neonate's skin and blood compartment sites (Fig. 3B,C). These data suggested that there was strong conservation of the viral population between mother and child following the perinatal transmission event. Finally, we detected high conservation of MVs between the neonate's skin and blood, suggesting a lack of compartmentalization—at least over this short time frame—that is distinct from what occurs during congenital HCMV transmission (Renzette et al. 2013). This suggests that HSV-1 did not undergo a bottleneck during dissemination from the skin into the blood, at least in this particular neonate.

Our findings could be influenced by the inferred high viral load present in the neonate following transmission, among the other unique factors associated with this particular case. With this data, we cannot determine whether this conclusion holds true for other routes of HSV-1 dissemination and/or viral spread (i.e., skin epithelium into the peripheral or central nervous system).

Several MVs were undetectable in the mother's HSV-1 genome, but appeared at detectable frequencies (de novo) in the neonate's viral genomes. One intriguing de novo MV was detected in the *UL6* gene of the HSV-1 genome derived from the neonate's skin (Table 1; Figs. 3C and 4). This gene encodes the portal protein, which forms a dodecameric ring at one apex of each procapsid and enables entry and packaging of newly synthesized viral DNA in a left-handed spool formation (Newcomb et al. 2001; Rochat et al. 2011; McElwee et al. 2018; Liu et al. 2019). The minor allele (thymidine) at this position causes a nonsynonymous mutation of glycine to tryptophan (G486W), a much bulkier and more hydrophobic residue with altered peptide backbone ϕ - ψ -torsion angles. There was extremely high AA conservation (99.3%) of UL6 across all 48 HSV-1 strains (data not shown) included in the phylogenetic network of Figure 6. However, the conservation of the portal protein was more varied when compared across 10 different species of the Simplexvirus genus (62.2% pairwise identity; Fig. 4B). The presence of this G486W variant at an AA position that was not well conserved across Simplexvirus species (Fig. 4B) suggested that this G486W change might be a functionally tolerable mutation for the virus present in the neonate's skin. If the low-level *UL6* variant were to become fixed in the viral population, it could impact DNA packaging of the virus and subsequent pathogenesis or virulence. The de novo appearance of this MV in the neonate's virus population suggests that immune pressures—or lack thereof in the neonate—shape viral evolution within a host over time (Baz et al. 2006; Bull et al. 2012; Vega et al. 2014).

The *UL30* gene, which encodes the crucial viral DNA polymerase, contained two variants of interest that were detected in all three samples—a low-frequency MV frameshift, and a unique variation in the overall consensus-level AA sequence (Table 1; Fig. 5A). The viral DNA polymerase enzyme is 1235 amino acids in length, with eight conserved regions present in the catalytic domain of the protein (Andrei et al. 2000). Of these two sites, the low-frequency MV in *UL30* causes a frameshift in the thumb domain of the enzyme (A1109) (Fig. 5B,D,E). Based on literature from the related bacteriophage RB69 polymerase, the thumb domain is hypothesized to interact directly with the DNA duplex during viral replication and editing (Liu et al. 2006). A frameshift mutation at A1109 in the viral polymerase would result in a 126-AA alternative tail, potentially altering enzyme function and/or eliminating the UL42 processivity factor's binding domain at the extreme carboxyl terminus of UL30.

The second, consensus-level AA variant in UL30 was detected in all three viral genomes and has not been documented in any previously published work (Table 2). Specifically, this variant occurred at the amino terminus of the enzyme, within the amino-terminal domain (AAs 141–362) (Fig. 5A–C; Liu et al. 2006). This region from AA 201–298 in the amino-terminal domain encodes a $\beta\alpha\beta\beta\alpha\beta$ fold, which is predicted to be an RNA-binding motif that contains a cluster of positively charged arginine and aromatic tyrosine residues (Burd and Dreyfuss 1994; Liu et al. 2006). In the published *UL30* structure, L279R is present in an α helix that is buried within the hydrophobic core of the protein domain. The mutation to arginine at this position may be hypothesized to affect binding of RNA to the amino-terminal domain of the viral polymerase through interference with domain folding.

Reports in the literature suggest that the majority of pregnant women who are infected with HSV prior to labor do not experience severe symptoms themselves, and that the main risk lies in acquisition of the virus in the late stages of pregnancy, near the time of labor, in which the probability of virus transmission to the fetus is greatly increased (Brown et al.

1997; Stephenson-Famy and Gardella 2014). Transmission of HSV between hosts is not well-characterized at the molecular or virus population level. The presence of a bottleneck at the point of transmission has been identified in other viruses, including HIV and influenza A virus, but has not yet been investigated for HSV-1 (Fischer et al. 2010; Masharsky et al. 2010; Frise et al. 2016; Leonard et al. 2017). From this case study it can be concluded that in the presence of maternal viremia, transmission of perinatal HSV-1 resulted in the transfer of the majority of the maternal viral population to the neonate. With no vaccine currently available for this virus and maternal–neonate transmission frequently occurring without maternal awareness of infection, an improved understanding of viral transmission events between human hosts is more critical than ever (Saral et al. 1981; Elion 1982; Gnann et al. 1983; Kimberlin 2001; Corey et al. 2004).

METHODS

Clinical Specimens

Because of the patient deaths involved in the case described here, the hospital laboratory decided to pursue complete viral genomic sequencing to determine if there were unexpected mutations in the HSV-1 samples from these individuals. Specimens from the pathology laboratory included already-isolated DNA from the mother's blood, already-isolated DNA from the neonate's skin swab, and a neonatal blood sample. These samples were the remainder of the specimens used in the diagnostic process. Both the isolated DNA samples were used directly for sample library prep and Illumina NGS, following HSV-specific quantitative real-time PCR (qPCR) and total DNA quantitation by Qubit 2.0 fluorimeter (Life Technologies-Invitrogen) at Pennsylvania State University. DNA was extracted from the neonate's blood for downstream library prep and NGS as detailed below.

DNA Extraction and HSV qPCR

DNA was extracted from the neonate's blood sample using phenol:chloroform:isoamyl alcohol in a ratio of 25:24:1 as previously described (Sambrook and Russell 2001). Extracted DNA was quantified using a Qubit 2.0 fluorimeter. Inert carrier in the form of linear polyacrylamide (10 µg) was added to each sample. Viral DNA was quantified using qPCR to detect the type-common region of HSV glycoprotein B gene *UL27* as previously described (Ryncarz et al. 1999; Johnston et al. 2014). This value was used to infer the number of gB copies and, by inference, the number of HSV genome copies in each sample.

Sample Library Prep, Illumina Sequencing, and Viral Assembly

Total DNA was sheared into ~800-base pair (bp) fragments using a Covaris sonicator with settings as follows: 10% duty, power 60, 200 cycles/burst for 60 sec at 4°C. Sheared DNA was processed using the KAPA LTP Library Preparation Kit compatible with Illumina platform reagents (KK8230). After ligation of SeqCap library adapters, a single postligation cleanup was performed before library DNA was amplified in the precapture ligation-mediated PCR (LM-PCR) step. The bait library was a custom Roche/Nimblegen SeqCap target enrichment platform designed and validated by the Bloom laboratory (A Dhummakupt, A Morse, and J Fear, unpubl.). The bait library included custom biotinylated DNA fragments that bind specifically to the DNA of HSV-1 strain 17 (GenBank JN555585.1), as well as to sequences of VZV strain Dumas (NC_001348.1), Epstein–Barr virus strain Akata (KC207813.1), and mouse gammaherpesvirus-68 strain WUMS (U97553.2). The probes ranged in length from 50 to 100 bp, with an average length of 75 bp. The depth of coverage for the targets was uniform across the entire capture space, with ~30% overlap. As part of the oligo-bait design process, any

proposed sequences with homology with the mouse, human, or rabbit genomes were excluded. Each sample library was hybridized for 72 h at 47°C with the oligonucleotide probes, after which the baits and attached captured viral DNA were isolated using magnetic beads. Following enrichment, a postcapture PCR step was performed before sequencing each bar-coded library on an Illumina MiSeq (v.3 chemistry) to obtain 300-bp paired-end sequence reads.

Prior to genome assembly, the sequence reads were passed through a series of quality control filters to remove Illumina primers and adapters and to trim low-quality terminal bases from each read (Szpara et al. 2010, 2011). A BLAST database consisting of all known HSV-1 and HSV-2 genomes was constructed, and sequence reads with an e -value $< 10^{-2}$ were used to build consensus genomes. Consensus HSV-1 genomes were obtained using a de novo viral genome assembly (VirGA) pipeline (Parsons et al. 2015). Briefly, eight SSAKE de novo assemblies were generated using varied parameters and combined into a draft genome for each sample using Celera and GapFiller (Myers et al. 2000; Warren et al. 2007; Boetzer and Pirovano 2012). We transferred annotations from the HSV-1 reference genome (strain 17; GenBank ID JN555585) to each new genome based on sequence homology. Sequencing statistics for the three clinical genomes generated in this work are listed in Supplemental Table S1.

Consensus Genome Comparison and Phylogenetic Analysis

Trimmed versions of viral genomes (lacking the terminal repeats) were used for consensus genome comparison, to avoid overrepresentation of the internal and terminal repeats (see Fig. 2A for illustration) (Szpara et al. 2014). MAFFT was used to construct pairwise global nucleotide alignments between whole-genome sequences (Kato et al. 2002). ClustalW2 was used for pairwise AA alignments between open reading frames (Larkin et al. 2007). Custom Python scripts were used to calculate protein-coding differences and DNA variation between samples. Phylogenetic networks were constructed with SplitsTree4, using the uncorrected P distance and ignoring all gaps (Huson 1998). The identity graph shown in Figure 2 was generated using Geneious software R11 (version 11.03) (Kearse et al. 2012). See Supplemental Table S3 for GenBank IDs and references for HSV-1 isolates used to construct the phylogenetic network.

Intrastrain Minor Variant Detection and Validation

Each consensus genome was analyzed for the presence of MV loci including single-nucleotide polymorphisms (SNPs) and indels. VarScan v2.2.11 was used to detect variants and indels in each consensus genome (Koboldt et al. 2012). Multiple parameters were applied to differentiate true MVs (SNPs and indels) from technical artifacts (Depledge et al. 2014): minimum variant allele frequency $\geq 2\%$ (0.02); base call quality ≥ 20 ; read depth at the position ≥ 100 ; independent reads supporting minor allele ≥ 5 . Polymorphisms with directional strand bias of $\geq 90\%$ were excluded. As noted in the text, the minimum variant allele frequency was set to a lower threshold of 0.2% (0.002) for selected analyses. SnpEff and SnpSift were used to map the MVs back to the consensus genomes and identify their distribution and potential mutational impact (Cingolani et al. 2012a, 2012b). All MVs were visually inspected using the Integrative Genomics Viewer (IGV) v2.3.97 to verify raw sequence read support for the VarScan output, and only MVs of high confidence were included and depicted in Figure 3. These data excluded clustered MVs with waning coverage ($< 1000\times$), typically found in highly repetitive regions of the genome. Of the MV calls included in Figure 3, 85% had sequence read coverage of $> 500\times$, with an average coverage of $> 2000\times$. See Supplemental Table S2 and ScholarSphere site (<https://doi.org/10.18113/S15M08>) for all MV data.

For transmission analysis of MVs from mother to child, we used a “donor” reference genome to track MVs from the donor virus population to the recipient’s virus population using a minimum variant allele frequency of $\geq 2\%$ (0.02) (Fig. 3B), akin to the approach used by Hage et al. (2017) for longitudinal samples from a single individual. For transmitted variants, the mother’s blood-derived viral genome was used as the reference or “donor” genome to detect MVs in both the neonate’s skin- and blood-derived viral genome populations, respectively (Fig. 3B). Analysis of disseminated MVs in the neonate was performed using the neonate’s skin-derived viral genome as the reference “donor” genome to detect MVs in the neonate’s blood-derived viral genome population (Fig. 3B).

To determine how conserved the UL6 protein was across a subset of nine Simplexvirus species, protein sequences were obtained from GenBank, and an AA alignment was created in Geneious software R11 (version 11.03) (Kearse et al. 2012). The AA alignment was visualized using Jalview v2.10.3b1 (Fig. 4B; Waterhouse et al. 2009).

Crystal Structure Modeling and Visualization

The crystal structure of the viral polymerase enzyme (UL30) from HSV-1 strain KOS (Liu et al. 2006) was visualized and manipulated using the PyMOL Molecular Graphics System, version 2.1.1, Schrödinger, LLC. In addition to the unique AA residue (L279R) shown in Figure 5 and highlighted in Tables 1 and 2, the UL30 protein from the mother’s and neonate’s HSV-1 genomes differs from that of the UL30 protein from KOS at A566T, in which the Ala is shared with 45 out of 48 other strains listed in Supplemental Table S3, and at V905M, where 11 other strains also have Val present at this position.

ADDITIONAL INFORMATION

Data Deposition and Access

GenBank accession numbers for the three viral genomes are as follows: Mother viremic blood HSV-1, MK952185; Neonate disseminated skin HSV-1, MK952183; and Neonate disseminated blood HSV-1, MK952184. Genomes, annotations, and Supplemental tables for the three HSV-1 genomes presented in this work are also archived in a ScholarSphere data repository: <https://doi.org/10.18113/S15M08>.

Ethics Statement

This case report was approved by the University of Iowa Institutional Review Board (No. 201903815).

Acknowledgments

The authors acknowledge support for the project’s intellectual development from members of the Szpara lab and colleagues.

Author Contributions

M.M.S. isolated viral DNA and prepared samples for targeted genomic enrichment and NGS. D.W.R. assembled full-length viral genomes and performed computational analyses. U.P. performed MV analyses. M.M.S. analyzed computational data output and generated all figures. D.C.B. contributed HSV oligonucleotide probes for targeted genomic enrichment. C.G., B.F., and M.L.S. conceptualized this work after a systematic review of the case reports. M.M.S. and M.L.S. wrote the manuscript, and all authors contributed to its completion.

Competing Interest Statement

The authors have declared no competing interest.

Received May 16, 2019; accepted in revised form September 4, 2019.

Funding

This project was funded, in part, by support from the Eberly College of Science and the Huck Institutes of the Life Sciences at Pennsylvania State University, as well as the National Institutes of Health (NIH) grant R21 AI140443 (M.L.S., PI; C.G. Contributor). This project was also funded, in part, under a grant with the Pennsylvania Department of Health using Commonwealth Universal Research Enhancement Program (CURE) funds (M.L.S.). D.C.B. acknowledges funding for the oligonucleotide enrichment probes from R01 AI048633.

REFERENCES

- Akhtar LN, Bowen CD, Renner DW, Pandey U, Della Fera AN, Kimberlin DW, Prichard MN, Whitley RJ, Weitzman MD, Szpara ML. 2019. Genotypic and phenotypic diversity of herpes simplex virus 2 within the infected neonatal population. *mSphere* **4**: e00590-18. doi:10.1128/mSphere.00590-18
- Altschul SF. 1991. Amino acid substitution matrices from an information theoretic perspective. *J Mol Biol* **219**: 555–565. doi:10.1016/0022-2836(91)90193-A
- Andrei G, Snoeck R, Clercq ED, Esnouf R, Fiten P, Opdenakker G. 2000. Resistance of herpes simplex virus type 1 against different phosphonylmethoxyalkyl derivatives of purines and pyrimidines due to specific mutations in the viral DNA polymerase gene. *J Gen Virol* **81**: 639–648. doi:10.1099/0022-1317-81-3-639
- Atanasiu D, Saw WT, Cohen GH, Eisenberg RJ. 2010. Cascade of events governing cell–cell fusion induced by herpes simplex virus glycoproteins gD, gH/gL, and gB. *J Virol* **84**: 12292–12299. doi:10.1128/JVI.01700-10
- Baz M, Abed Y, McDonald J, Boivin G. 2006. Characterization of multidrug-resistant influenza A/H3N2 viruses shed during 1 year by an immunocompromised child. *Clin Infect Dis* **43**: 1555–1561. doi:10.1086/508777
- Bell S, Cranage M, Borysiewicz L, Minson T. 1990. Induction of immunoglobulin G Fc receptors by recombinant vaccinia viruses expressing glycoproteins E and I of herpes simplex virus type 1. *J Virol* **64**: 2181–2186.
- Berrington WR, Jerome KR, Cook L, Wald A, Corey L, Casper C. 2009. Clinical correlates of herpes simplex virus viremia among hospitalized adults. *Clin Infect Dis* **49**: 1295–1301. doi:10.1086/606053
- Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol* **13**: 1–9. doi:10.1186/gb-2012-13-6-r56
- Bordería AV, Isakov O, Moratorio G, Henningsson R, Agüera-González S, Organtini L, Gnädig NF, Blanc H, Alcover A, Hafenstein S, et al. 2015. Group selection and contribution of minority variants during virus adaptation determines virus fitness and phenotype. *PLoS Pathog* **11**: e1004838. doi:10.1371/journal.ppat.1004838
- Brown ZA, Selke S, Zeh J, Kopelman J, Maslow A, Ashley RL, Watts DH, Berry S, Herd M, Corey L. 1997. The acquisition of herpes simplex virus during pregnancy. *N Engl J Med* **337**: 509–516. doi:10.1056/NEJM199708213370801
- Brown ZA, Wald A, Morrow RA, Selke S, Zeh J, Corey L. 2003. Effect of serologic status and cesarean delivery on transmission rates of herpes simplex virus from mother to infant. *J Am Med Assoc* **289**: 203–209. doi:10.1001/jama.289.2.203
- Bull RA, Eden J-S, Luciani F, McElroy K, Rawlinson WD, White PA. 2012. Contribution of intra- and interhost dynamics to norovirus evolution. *J Virol* **86**: 3219–3229. doi:10.1128/JVI.06712-11
- Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**: 615–621. doi:10.1126/science.8036511
- Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X. 2012a. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front Genet* **3**: 35. doi:10.3389/fgene.2012.00035
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012b. A program for annotating and predicting the effects of single nucleotide polymorphisms. *SnpEff. Fly (Austin)* **6**: 80–92. doi:10.4161/fly.19695
- Corey L, Wald A. 2009. Maternal and neonatal HSV infections. *N Engl J Med* **361**: 1376–1385. doi:10.1056/NEJMra0807633
- Corey L, Wald A, Patel R, Sacks SL, Tying SK, Warren T, Douglas Jr JM, Paavonen J, Morrow RA, Beutner KR, et al. 2004. Once-daily valacyclovir to reduce the risk of transmission of genital herpes. *N Engl J Med* **350**: 11–20. doi:10.1056/NEJMoa035144
- Crute J, Lehman IR. 1989. Herpes simplex-1 DNA polymerase. *J Biol Chem* **254**: 19266–19270.
- Depledge DP, Kundu S, Jensen NJ, Gray ER, Jones M, Steinberg S, Gershon A, Kinchington PR, Schmid DS, Balloux F, et al. 2014. 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. doi:10.1093/molbev/mst210
- Depledge DP, Brown J, Macanovic J, Underhill G, Breuer J. 2016. Viral genome sequencing proves nosocomial transmission of fatal varicella. *J Infect Dis* **214**: 1399–1402. doi:10.1093/infdis/jiw398
- Dingwell KS, Brunetti CR, Hendricks RL, Tang Q, Tang M, Rainbow AJ, Johnson DC. 1994. Herpes simplex virus glycoproteins E and I facilitate cell-to-cell spread in vivo and across junctions of cultured cells. *J Virol* **68**: 834–845.
- Dingwell KS, Doering LC, Johnson DC. 1995. Glycoproteins E and I facilitate neuron-to-neuron spread of herpes simplex virus. *J Virol* **69**: 7087–7098.
- Drake JW, Hwang CBC. 2005. On the mutation rate of herpes simplex virus type 1. *Genetics* **170**: 969–970. doi:10.1534/genetics.104.040410
- Eliou GB. 1982. Mechanism of action and selectivity of acyclovir. *Am J Med* **73**: 7–13. doi:10.1016/0002-9343(82)90055-9
- Farooq AV, Shukla D. 2012. Herpes simplex epithelial and stromal keratitis: an epidemiologic update. *Surv Ophthalmol* **57**: 448–462. doi:10.1016/j.survophthal.2012.01.005
- Firth C, Kitchen A, Shapiro B, Suchard MA, Holmes EC, Rambaut A. 2010. Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol Biol Evol* **27**: 2038–2051. doi:10.1093/molbev/msq088
- Fischer W, Ganusov VV, Giorgi EE, Hraber PT, Keele BF, Leitner T, Han CS, Gleasner CD, Green L, Lo C-C, et al. 2010. Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS One* **5**: e12303. doi:10.1371/journal.pone.0012303
- Frise R, Bradley K, van Doremalen N, Galiano M, Elderfield RA, Stilwell P, Ashcroft JW, Fernandez-Alonso M, Miah S, Lackenby A, et al. 2016. Contact transmission of influenza virus between ferrets imposes a looser bottleneck than respiratory droplet transmission allowing propagation of antiviral resistance. *Sci Rep* **6**: 29793. doi:10.1038/srep29793
- Gnann JW, Barton NH, Whitley RJ. 1983. Acyclovir: mechanism of action, pharmacokinetics, safety and clinical applications. *Pharmacother J Hum Pharmacol Drug Ther* **3**: 275–283. doi:10.1002/j.1875-9114.1983.tb03274.x
- Gompels UA. 1991. Characterization and sequence analyses of antibody-selected antigenic variants of herpes simplex virus show a conformationally complex epitope on glycoprotein H. *J Virol* **65**: 9.
- Grose C. 1994. Congenital infections caused by varicella zoster virus and herpes simplex virus. *Semin Pediatr Neurol* **1**: 43–49.
- Grose C, Tyler S, Peters G, Hiebert J, Stephens GM, Ruyechan WT, Jackson W, Storlie J, Tipples GA. 2004. Complete DNA sequence analyses of the first two varicella-zoster virus glycoprotein E (D150N) mutant viruses found in North America: evolution of genotypes with an accelerated cell spread phenotype. *J Virol* **78**: 6799–6807. doi:10.1128/JVI.78.13.6799-6807.2004
- Hage E, Wilkie GS, Linnenweber-Held S, Dhingra A, Suárez NM, Schmidt JJ, Kay-Fedorov PC, Mischak-Weissinger E, Heim A, Schwarz A, et al. 2017. Characterization of human cytomegalovirus genome diversity in immunocompromised hosts by whole-genome sequencing directly from clinical specimens. *J Infect Dis* **215**: 1673–1683. doi:10.1093/infdis/jix157
- Hall JD, Almy RE. 1982. Evidence for control of herpes simplex virus mutagenesis by the viral DNA polymerase. *Virology* **116**: 535–543. doi:10.1016/0042-6822(82)90146-5
- Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci* **89**: 10915–10919. doi:10.1073/pnas.89.22.10915
- Huson DH. 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**: 68–73. doi:10.1093/bioinformatics/14.1.68
- Hutchinson L, Browne H, Wargent V, Davis-Poynter N, Primorac S. 1992. A novel herpes simplex virus glycoprotein, gL, forms a complex with glycoprotein H (gH) and affects normal folding and surface expression of gH. *J Virol* **66**: 11.
- Johnston C, Magaret A, Selke S, Remington M, Corey L, Wald A. 2008. Herpes simplex virus viremia during primary genital infection. *J Infect Dis* **198**: 31–34. doi:10.1086/588676
- Johnston C, Zhu J, Jing L, Laing KJ, McClurkan CM, Klock A, Diem K, Jin L, Stanaway J, Tronstein E, et al. 2014. Virologic and immunologic evidence of multifocal genital herpes simplex virus 2 infection. *J Virol* **88**: 4921–4931. doi:10.1128/JVI.03285-13
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059–3066. doi:10.1093/nar/gkf436
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649. doi:10.1093/bioinformatics/bts199
- Kimberlin DW. 2001. Acyclovir derivatives and other new antiviral agents. *Semin Pediatr Infect Dis* **12**: 224–234. doi:10.1053/spid.2001.24097

- Kimberlin DW. 2007. Herpes simplex virus infections of the newborn. *Semin Perinatol* **31**: 19–25. doi:10.1053/j.semperi.2007.01.003
- Kimberlin DW, Lin CY, Jacobs RF, Powell DA, Corey L, Gruber WC, Rathore M, Bradley JS, Diaz PS, Kumar M, et al. 2001a. Safety and efficacy of high-dose intravenous acyclovir in the management of neonatal herpes simplex virus infections. *Pediatrics* **108**: 230–238. doi:10.1542/peds.108.2.230
- Kimberlin DW, Lin CY, Jacobs RF, Powell DA, Frenkel LM, Gruber WC, Rathore M, Bradley JS, Diaz PS, Kumar M, et al. 2001b. Natural history of neonatal herpes simplex virus infections in the acyclovir era. *Pediatrics* **108**: 223–229. doi:10.1542/peds.108.2.223
- Kimura H, Futamura M, Kito H, Ando T, Goto M, Kuzushima K, Shibata M, Morishima T. 1991. Detection of viral DNA in neonatal herpes simplex virus infections: frequent and prolonged presence in serum and cerebrospinal fluid. *J Infect Dis* **164**: 289–293. doi:10.1093/infdis/164.2.289
- Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22**: 568–576. doi:10.1101/gr.129684.111
- 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. doi:10.1371/journal.ppat.1005499
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinforma Appl Note* **23**: 2947–2948. doi:10.1093/bioinformatics/btm404
- Leonard AS, Weissman DB, Greenbaum B, Ghedin E, Koelle K. 2017. Transmission bottleneck size estimation from pathogen deep-sequencing data, with an application to human influenza A virus. *J Virol* **91**: 1–19. doi:10.1101/101790
- Liu S, Knafels JD, Chang JS, Waszak GA, Baldwin ET, Deibel MR, Thomsen DR, Homa FL, Wells PA, Tory MC, et al. 2006. Crystal structure of the herpes simplex virus 1 DNA polymerase. *J Biol Chem* **281**: 18193–18200. doi:10.1074/jbc.M602414200
- Liu Y-T, Jih J, Dai X, Bi G-Q, Zhou ZH. 2019. Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. *Nature* **570**: 257–261. doi:10.1038/s41586-019-1248-6
- Looker KJ, Magaret AS, May MT, Turner KM, Vickerman P, Newman LM, Gottlieb SL. 2017. First estimates of the global and regional incidence of neonatal herpes infection. *Lancet Glob Health* **5**: e300–e309. doi:10.1016/S2214-109X(16)30362-X
- Looker KJ, Magaret AS, May MT, Turner KME, Vickerman P, Gottlieb SL, Newman LM. 2015a. Global and regional estimates of prevalent and incident herpes simplex virus type 1 infections in 2012. *PLoS One* **10**: e0140765. doi:10.1371/journal.pone.0140765
- Looker KJ, Magaret AS, Turner KME, Vickerman P, Gottlieb SL, Newman LM. 2015b. Global estimates of prevalent and incident herpes simplex virus type 2 infections in 2012. *PLoS One* **10**: e114989. doi:10.1371/journal.pone.0114989
- Mapelli M, Panjekar S, Tucker PA. 2005. The crystal structure of the herpes simplex virus 1 ssDNA-binding protein suggests the structural basis for flexible, cooperative single-stranded DNA binding. *J Biol Chem* **280**: 2990–2997. doi:10.1074/jbc.M406780200
- Masharsky AE, Dukhovlinova EN, Verevochkin SV, Tousseva OV, Skochilov RV, Anderson JA, Hoffman I, Cohen MS, Swanstrom R, Kozlov AP. 2010. A substantial transmission bottleneck among newly and recently HIV-1-infected injection drug users in St Petersburg, Russia. *J Infect Dis* **201**: 1697–1702. doi:10.1086/652702
- McElwee M, Vijayakrishnan S, Rixon F, Bhella D. 2018. Structure of the herpes simplex virus portal-vertex. *PLoS Biol* **16**: e2006191. doi:10.1371/journal.pbio.2006191
- McGeoch DJ, Dolan A, Donald S, Street C. 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. *Nucleic Acids Res* **14**: 1727–1746. doi:10.1093/nar/14.4.1727
- McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC, McNab D, Perry LJ, Scott JE, Taylor P. 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J Gen Virol* **69**: 1531–1574. doi:10.1099/0022-1317-69-7-1531
- Myers EW, Sutton G, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KHJ, Remington KA, et al. 2000. A whole-genome assembly of *Drosophila*. *Science* **287**: 2196–2204. doi:10.1126/science.287.5461.2196
- Nellisery JK, Szczepaniak R, Lamberti C, Weller SK. 2007. A putative leucine zipper within the herpes simplex virus type 1 UL6 protein is required for portal ring formation. *J Virol* **81**: 8868–8877. doi:10.1128/JVI.00739-07
- Newcomb WW, Juhas RM, Thomsen DR, Homa FL, Burch AD, Weller SK, Brown JC. 2001. The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *J Virol* **75**: 10923–10932. doi:10.1128/JVI.75.22.10923-10932.2001

- Pandey U, Renner DW, Thompson RL, Szpara ML, Sawtell NM. 2017. Inferred father-to-son transmission of herpes simplex virus results in near-perfect preservation of viral genome identity and in vivo phenotypes. *Sci Rep* **7**: 13666. doi:10.1038/s41598-017-13936-6
- 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. doi:10.1128/mBio.02213-14
- Peng T, Ponce de Leon M, Novotny MJ, Jiang H, Lambris JD, Dubin G, Spear PG, Cohen GH, Eisenberg RJ. 1998. Structural and antigenic analysis of a truncated form of the herpes simplex virus glycoprotein gH-gL complex. *J Virol* **72**: 6092–6103.
- Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z, et al. 2012. ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res* **40**: D593–D598. doi:10.1093/nar/gkr859
- Renzette N, Gibson L, Bhattacharjee B, Fisher D, Schleiss MR, Jensen JD, Kowalik TF. 2013. Rapid intrahost evolution of human cytomegalovirus is shaped by demography and positive selection. *PLoS Genet* **9**: 1–14. doi:10.1371/journal.pgen.1003735
- Rochat RH, Liu X, Murata K, Nagayama K, Rixon FJ, Chiu W. 2011. Seeing the portal in herpes simplex virus type 1 B capsids. *J Virol* **85**: 1871–1874. doi:10.1128/JVI.01663-10
- Roizman B, Knipe DM, Whitley R. 2013. Herpes simplex viruses. In *Fields virology* (ed. Knipe DM, Howley PM), Vol. 2, pp. 1823–1897. Lippincott Williams & Wilkins, Philadelphia.
- Ryncarz AJ, Goddard J, Wald A, Huang M-L, Roizman B, Corey L. 1999. Development of a high-throughput quantitative assay for detecting herpes simplex virus DNA in clinical samples. *J Clin Microbiol* **37**: 1941–1947.
- Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Santos RA, Padilla JA, Hatfield C, Grose C. 1998. Antigenic variation of varicella zoster virus fc receptor gE: loss of a major B cell epitope in the ectodomain. *Virology* **249**: 21–31. doi:10.1006/viro.1998.9313
- Santos RA, Hatfield CC, Cole NL, Padilla JA, Moffat JF, Arvin AM, Ruyechan WT, Hay J, Grose C. 2000. Varicella-zoster virus gE escape mutant VZV-MSP exhibits an accelerated cell-to-cell spread phenotype in both infected cell cultures and SCID-hu mice. *Virology* **275**: 306–317. doi:10.1006/viro.2000.0507
- Saral R, Burns WH, Laskin OL, Santos GW, Lietman PS. 1981. Acyclovir prophylaxis of herpes-simplex-virus infections. *N Engl J Med* **305**: 63–67. doi:10.1056/NEJM198107093050202
- Stephenson-Famy A, Gardella C. 2014. Herpes simplex virus infection during pregnancy. *Obstet Gynecol Clin North Am* **41**: 601–614. doi:10.1016/j.ogc.2014.08.006
- Streilein JW, Dana MR, Ksander BR. 1997. Immunity causing blindness: five different paths to herpes stromal keratitis. *Immunity Today* **18**: 443–449. doi:10.1016/S0167-5699(97)01114-6
- Szpara ML, Parsons L, Enquist LW. 2010. Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. *J Virol* **84**: 5303–5313. doi:10.1128/JVI.00312-10
- Szpara ML, Tafuri YR, Parsons L, Shamim SR, Verstrepen KJ, Legendre M, Enquist LW. 2011. A wide extent of inter-strain diversity in virulent and vaccine strains of alphaherpesviruses. *PLoS Pathog* **7**: 1–23. doi:10.1371/journal.ppat.1002282
- 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. doi:10.1128/JVI.01987-13
- Vega E, Donaldson E, Huynh J, Barclay L, Lopman B, Baric R, Chen LF, Vinje J. 2014. RNA populations in immunocompromised patients as reservoirs for novel norovirus variants. *J Virol* **88**: 14184–14196. doi:10.1128/JVI.02494-14
- Warren RL, Sutton GG, Jones SJM, Holt RA. 2007. Assembling millions of short DNA sequences using SSAKE. *Bioinformatics* **23**: 500–501. doi:10.1093/bioinformatics/btl629
- Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189–1191. doi:10.1093/bioinformatics/btp033
- Weller SK, Coen DM. 2012. Herpes simplex viruses: mechanisms of DNA replication. *Cold Spring Harb Perspect Biol* **4**: a013011. doi:10.1101/cshperspect.a013011
- Whitley RJ. 1986. Neonatal herpes simplex virus infections. Presentation and management. *J Reprod Med* **31**: 426–432.
- Whitley RJ. 2006. Herpes simplex encephalitis: adolescents and adults. *Antiviral Res* **71**: 141–148. doi:10.1016/j.antiviral.2006.04.002
- Williams LE, Nesburn AB, Kaufman HE. 1965. Experimental induction of disciform keratitis. *Arch Ophthalmol* **73**: 112–114. doi:10.1001/archopht.1965.00970030114023
- Wirgart BZ, Estrada V, Jackson W, Linde A, Grose C. 2006. A novel varicella-zoster virus gE mutation discovered in two Swedish isolates. *J Clin Virol* **37**: 134–136. doi:10.1016/j.jcv.2006.06.007