Genotypic and phenotypic diversity within the neonatal HSV-2 population

Lisa N. Akhtar[1], Christopher D. Bowen[2], Daniel W. Renner[2], Utsav Pandey[2], Ashley N. Della Fera[3], David W. Kimberlin[4], Mark N. Prichard[4], Richard J. Whitley[4], Matthew D. Weitzman[3,5]*, Moriah L. Szpara[2]*

[1] Department of Pediatrics, Division of Infectious Diseases, Children's Hospital of Philadelphia and University of Pennsylvania Perelman School of Medicine

[2] Department of Biochemistry and Molecular Biology, Center for Infectious Disease Dynamics, and the Huck Institutes of the Life Sciences, Pennsylvania State University

[3] Division of Protective Immunity and Division of Cancer Pathobiology, Children's Hospital of Philadelphia

[4] Department of Pediatrics, Division of Infectious Diseases, University of Alabama at Birmingham

[5] Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine

***Corresponding Authors**
Moriah L. Szpara
Dept. of Biochemistry & Molecular Biology
The Huck Institutes of the Life Sciences
W-208 Millennium Science Complex
(MSC)
Pennsylvania State University
University Park, PA 16802 USA
Phone: 814-867-0008
Email: moriah@psu.edu

Matthew D. Weitzman
Division of Protective Immunity
The Children's Hospital of Philadelphia
4050 Colket Translational Research
Building
3501 Civic Center Blvd
Philadelphia, PA 19104-4318
Phone: 267-425-2068
Email: weitzmanm@email.chop.edu

Author emails:
Lisa N. Akhtar: AkhtarL@email.chop.edu
Christopher D. Bowen: cdb5004@psu.edu
Daniel W. Renner: dwr19@psu.edu
Utsav Pandey: uup104@psu.edu
Ashley N. Della Fera: coxan@email.chop.edu
David W. Kimberlin: DKimberlin@peds.uab.edu
Mark N. Prichard: MPrichard@peds.uab.edu
Richard J. Whitley: RWhitley@peds.uab.edu
Matthew D. Weitzman: weitzmanm@email.chop.edu
Moriah L. Szpara: moriah@psu.edu

Keywords: Neonatal, herpes simplex virus 2, human herpesvirus 2, comparative genomics, viral spread

Abstract

Neonates infected with herpes simplex virus (HSV) at the time of birth can have different courses of clinical disease. Approximately half of those infected display manifestations limited to the skin, eyes, or mouth (SEM disease, 45%). However, others develop invasive infections that spread systemically (disseminated, 25%) or to the central nervous system (CNS, 30%); both of which are associated with significant morbidity and mortality. The viral and/or host factors that predispose a neonate to these invasive forms of HSV infection are not known. To define the level of viral diversity within the neonatal population we evaluated ten HSV-2 isolates cultured from neonates with a range of clinical presentations. To assess viral fitness independent of host immune factors, we measured viral growth characteristics of each isolate in cultured cells. We found that HSV-2 isolates displayed diverse *in vitro* phenotypes. Isolates from neonates with CNS disease were associated with larger average plaque size and enhanced spread through culture, with isolates derived directly from the cerebrospinal fluid (CSF) exhibiting the most robust growth characteristics. We then sequenced the complete viral genomes of all ten neonatal HSV-2 isolates, providing the first insights into HSV genomic diversity in this clinical setting. We found extensive inter-host variation between isolates distributed throughout the HSV-2

genome. Furthermore, we assessed intra-host variation and found that each HSV-2 isolate contained minority variants, with two viral isolates containing ten-fold higher levels of allelic variation than other neonatal isolates or comparable adult isolates. HSV-2 glycoprotein G (gG, US4), gI (US7), gK (UL53), and viral proteins UL8, UL20, UL24, and US2 contained variants that were found only in neonatal isolates associated with CNS disease. Many of these genes encode viral proteins known to contribute to cell-to-cell spread and/or neurovirulence in mouse models of CNS disease. This study represents the first-ever application of comparative pathogen genomics to neonatal HSV disease.

Introduction

More than 14,000 neonates are infected with HSV each year worldwide (1), including 1500 neonates in the United States. Infants are typically infected at the time of birth due to maternal genital shedding of HSV-1 or HSV-2, often by mothers who are not aware of their infection (2–4). The recent increase in genital HSV-1 incidence among women of childbearing age, particularly in developed nations, suggests that the burden of neonatal infection will continue to rise (1, 5). While some infected infants exhibit only superficial infection limited to the skin, eyes, or mouth (SEM disease), about half develop invasive systemic (disseminated disease) or central nervous system (CNS disease) infections associated with significant morbidity and mortality (6, 7). Currently, the antiviral medication acyclovir is the standard therapy for all forms of neonatal HSV disease. Although this intervention has reduced mortality due to invasive disease, most survivors are left with permanent neurodevelopmental deficits (8, 9).

The factors that predispose a neonate to invasive HSV infection are not known. Recent studies have attempted to address this question in adult patients, and shown that individuals who experience HSV infection of the brain often contain a host genetic defect within the Toll-like receptor-3 (TLR3) pathway (10, 11). The epidemiology of adult HSV encephalitis is consistent with this finding, in that both HSV infection of the CNS and the host TLR3 mutation are exceedingly rare events. By contrast, half of HSV-infected neonates experience invasive CNS or disseminated disease making it less likely that host genetic defects alone could account for invasive forms of infection. Instead, prior clinical data on mother-to-infant transmission of HSV indicate that most cases of neonatal disease result from newly acquired HSV infection with an absence of maternal immune protection (3, 4, 2). This suggests a window of opportunity where the contributions of viral genetic variation to the progression of infection and disease may be greater than in adults.

Prior studies have identified viral genetic factors that influence virulence or disease for reoviruses, influenza virus, HIV, and others (12–17). Recent data suggest that HSV harbors extensive diversity between strains (inter-strain variation) that may likewise influence the outcome(s) of infection, and that HSV genetic diversity may even exist within an infected individual (intra-strain variation) (18, 19). An early analysis by Rosenthal and colleagues provided proof of principle that a heterogeneous HSV population can exist in an invasive neonatal infection (20, 21). Recent advances in high-throughput sequencing (HTSeq) have now enabled similar studies to analyze herpesvirus genome-wide variation, and to detect the presence of minority variants within a single viral isolate or patient (22). These minor alleles can manifest as the new dominant allele or genotype after bottlenecks or selective pressure such as antiviral

therapy (23, 24). In HTSeq studies of congenital infections by the beta-herpesvirus human cytomegalovirus (HCMV), Renzette *et al.* found evidence for heterogeneous viral populations both within and between hosts (25–30), which far exceeded the levels of diversity observed in adult HCMV infections (31–33). Examination of vaccine-associated rashes for varicella zoster virus (VZV; an alpha-herpesvirus) has demonstrated that adult skin vesicles contained a subset of the viral population introduced during vaccination, and found at least 11 VZV genomic loci that were linked to rash formation (34). Together, these studies provide a precedent for viral genetic variation within and between hosts. These studies also demonstrate the potential of genome-wide analyses to illuminate viral genetic loci that may contribute to disease manifestations, particularly in the setting of the limited neonatal immune response.

Until recently, several technical barriers prevented thorough assessment of neonatal HSV genomes. A key constraint on studies of neonatal disease has been the availability of viral isolates associated with clinical information that have been maintained in a state sufficient for expansion in culture and sequencing. Specifically, the routine culturing of HSV is no longer part of the standard diagnostic workflow as it has been superseded by the speed and sensitivity of viral detection by PCR, limiting the number of cultured samples that are currently collected. Many previously archived neonatal isolates have been passaged extensively, allowing them to acquire mutations associated with adaptation to culture (20, 21). Additionally, the HSV genome is challenging to sequence and assemble with high confidence, due to its large size (~152 kb), high G+C content (~70%), and repeat-rich genome (22). Many earlier studies of HSV diversity (35–39) or the effect of HSV genetic variation on disease (40–42) have relied on low-resolution restriction fragment length polymorphism (RFLP) or single-gene PCR analyses.

Here we analyzed a set of 10 low-passage HSV-2 isolates collected from neonates with culture- or PCR-confirmed HSV infection, enrolled in three published clinical studies (8, 43, 44). These samples represented a wide range of clinical manifestations including SEM, CNS, and disseminated disease, and were associated with robust de-identified clinical information. We defined the level of diversity in this population using comparative genomics and an array of cell-based phenotypic assays. We found that HSV-2 isolates display diverse *in vitro* phenotypes, with the isolates derived from neonates with CNS disease displaying the largest average plaque size and enhanced spread through culture. We found extensive inter- and intra-host diversity distributed throughout the HSV-2 genome. Finally, we found coding variations in several HSV-2 proteins unique to isolates from neonates with CNS disease. Many of these viral proteins are known to modulate cell-to-cell spread and/or contribute to neurovirulence in mouse models of CNS disease. This study represents the first-ever application of comparative pathogen genomics to neonatal HSV disease and provides a basis for further exploration of genotype-phenotype links in this clinically vulnerable patient population.

Results

Neonatal HSV-2 samples represent a diverse clinical population

We utilized samples collected from ten HSV-2-infected neonates enrolled by the National Institute of Allergy and Infectious Diseases Collaborative Antiviral Study Group (CASG) for clinical trials between 1981 and 2008 (8, 43, 44). These infants encompassed a range of clinical disease manifestations (see **Table 1**), with about half experiencing invasive CNS disease (5 patients) or disseminated (DISS) disease with CNS involvement (2 patients), and the remainder

experiencing non-invasive SEM disease (3 patients). Extensive clinical information is available for each patient, including long-term neurocognitive and motor outcome (**Table 1**). This population is also diverse with respect to sex, race, gestational age, and enrollment center (**Table 1**; enrollment center data not shown). All samples were collected at the time of diagnosis, prior to initiation of acyclovir therapy. Each isolate was cultured once as part of the diagnostic process with expansion only for the experiments shown here. Although this is a relatively small group of samples due to the constraints discussed previously, it is the largest group of neonatal HSV samples ever compared genotypically or phenotypically, and is similar in size to prior studies of HCMV-infected neonates (25–30).

Neonatal HSV-2 isolates have different fitness in culture

To determine whether the viruses isolated from this neonatal population **(Table 1)** were intrinsically different, we first assessed viral growth in culture independent of the neonatal host and immune response. To minimize the impact of immune regulation we selected Vero monkey kidney cells, due to their lack of an interferon response (45, 46). Each viral isolate was applied to a confluent monolayer of cells *in vitro*, and allowed to form plaques for 100 hours (**Figure 1A**). Average plaque size differed between isolates, with large plaques being more frequent among viruses causing neonatal CNS disease (**Figure 1B**). Viruses isolated directly from the cerebrospinal fluid (CSF) (CNS11 and DISS14) produced plaques with an average size that was statistically larger than those isolated from the skin. Plaque size was assessed throughout passage in culture and remained constant from the time the isolates were received in our laboratory (passage 2) through their genetic and phenotypic analysis (passage 4). The low-passage HSV-2 clinical isolate SD90e, which was isolated from an adult patient, was used as a control for

comparison (47). The distribution or variance in plaque sizes produced by a given isolate is not statistically different between isolates. These differences in average plaque size between isolates in culture suggested that the HSV-2 populations found in each neonatal isolate are intrinsically different.

Entry kinetics, DNA replication, protein expression, and virus production do not account for differences in plaque size

Plaque formation is a complex endpoint that involves the ability of the virus to enter the cell, replicate its double stranded DNA genome, produce viral proteins and assemble new virions that spread to adjacent cells. We therefore explored whether the differences in plaque formation might reflect inherent differences in viral entry, DNA replication, protein production, or infectious virus production for neonatal HSV-2 isolates. Using the representative subset of four neonatal HSV isolates, we compared their rate of cell entry in an assay involving application of virus to chilled cells, followed by warming to synchronize cell entry **(Figure 2A).** A low pH solution was applied at various points over the first hour of cell entry to inactivate any virus that had not yet entered a cell. We found no difference in rates of cell entry between these four representative viral isolates (**Figure 2B**), suggesting that large plaques do not result from an increased rate of virus entry into cells. Likewise, over the time-course of a high MOI infection (MOI=5), we found that all four isolates produced a similar number of genome copies (as measured by qPCR for the gB gene; see Methods for details) **(Figure 2C),** and a similar number of plaque-forming units (**Figure 2D, E**). In order to rule out the possibility that very small foci of infection might be missed for the small plaque-forming isolates, we tested each isolate for its ability to form plaques on the highly-permissive U2OS human bone osteosarcoma epithelial cell

line. These cells lack innate sensing of viral infection through the STING pathway (48) and can even support the growth of highly-defective HSV-isolates that lack ICP0 function (49). All isolates were capable of forming large plaques on U2OS cells by 100 hpi, a time point at which half of the isolates still displayed a small-plaque phenotype on Vero cells (**Supplemental Figure 1**). We further quantified the production of infectious virus by counting plaque-forming-units from a single round of infection (a single-step growth curve) on both Vero and U2OS **(Figure 2D-2E)** cells in parallel; this likewise revealed no differences in virus production. Finally, we compared viral protein production for these isolates, and found no differences in the expression levels of a panel of viral proteins at either early (6 hpi) or late (24 hpi) time points of a single round (MOI = 5) of infection (**Figure 2F-2G**). These results suggest that large- and small-plaque-forming neonatal HSV-2 isolates do not differ significantly in viral entry, DNA replication, protein production, or infectious virus production over a single round of infection.

A subset of isolates exhibit enhanced cell-to-cell spread

In the absence of an alternative rationale for the ability of certain neonatal HSV-2 isolates to form larger plaques in culture, we assessed the ability of representative large- and small-plaque-forming isolates to spread from cell-to-cell. Vero cell monolayers were infected at low MOI (MOI = 0.001) in order to assess differences over multiple rounds of viral replication and spread throughout the cell monolayer over time (**Figure 3A**). The contribution of indirect cell-free spread was minimized by including 0.1% human serum in the media, and changing the media every 24 hours to remove released virus and refresh serum levels. Over a 72-hour time course, cell monolayers were assessed for the extent of cell-to-cell spread, either by harvesting cells and evaluating the quantity of virus by titer on U2OS (**Figure 3B**) or Vero **(Figure 3C)** cells, or by

fixing infected cell monolayers and evaluating the distribution of virus by immunofluorescence (**Figure 3D**, **3E**, and **Supplemental Figure 2**). Viral titers recovered from harvested cells were similar at 2 hpi, confirming that equivalent amounts of each virus were present following initial infection (**Figure 3B**). However, by 72 hpi, the large-plaque-forming isolates CNS11 and CNS03 had achieved viral titers significantly greater than those of the small-plaque-forming isolates CNS12 and SEM02 (**Figure 3B**).

We also directly evaluated cell-to-cell spread by immunostaining and quantifying the distribution of infected cells around each infectious focus. Infected Vero cell monolayers were fixed at 24, 48, and 72 hpi, and subjected to fluorescent immunocytochemistry with an antibody directed against total HSV (**Figure 3D** and **Supplemental Figure 2**). The region of HSV-positive cells surrounding a single initial infection is greater for the large-plaque-forming neonatal HSV-2 isolates at 24 hpi, and increases by 48 and 72 hpi. The central cytolytic clearings seen in the CNS11- and CNS03-infected monolayers were approximately 2-fold greater than those for CNS12 and SEM02, reflecting the average 2-fold increase in plaque size observed in the methylene-blue plaque staining observed in **Figure 1**. However, the region of infected cells surrounding the central cytolytic clearing for CNS11 and CNS03 was dramatically larger than for CNS12 and SEM02, suggesting that large-plaque-forming neonatal HSV-2 isolates have an even greater rate of spread from cell-to-cell than that observed by measuring plaque size alone. To quantify this increase in the distribution of infected cells after a low MOI infection, each coverslip was imaged **(Supplemental Figure 2)** and the total number of immunofluorescent pixels was quantified (**Figure 3E**). By 72 hpi, the area of HSV-infected cells was statistically greater for the large-plaque-forming isolates CNS11 and CNS03, as compared to the small-

plaque-forming isolates CNS12 and SEM02. Together these data indicated that large-plaque-forming isolates associated with CNS disease shared an enhanced ability to spread cell-to-cell through culture, in comparison to small-plaque-forming isolates.

Comparative genomics reveals genetic diversity in neonatal HSV-2 isolates

The differences identified in cell-to-cell spread between neonatal isolates in culture indicated intrinsic differences between these viruses. To reveal how genetic variation may contribute to culture phenotype, and ultimately clinical disease manifestations, we sequenced the complete viral genome for all ten neonatal HSV-2 isolates. The clinical trials utilized in this study enrolled HSV-infected infants from multiple enrollment sites across the United States (8, 43, 44). Therefore, we first assessed the overall degree of relatedness between these isolates to understand whether any similarities in viral or geographic origin might have contributed to *in vitro* or clinical phenotype. In light of the known propensity for a history of recombination between HSV strains, we used a graph-based network to investigate the phylogenetic relationship between these isolates. We found a similar degree of divergence among all ten HSV-2 isolates (**Figure 4A**), which was further corroborated by their dispersion in a network graph of all available HSV-2 genomes from GenBank (**Fig. 4B**, see **Supplemental Table 1** for HSV-2 genome list). This suggested that similarities in viral genetic origin were not responsible for determining the clinical manifestations of neonatal HSV-2 infection.

Overall protein-coding diversity in neonatal HSV-2 isolates is similar to that observed in adult HSV-2 strains

We next asked whether overt defects in any single HSV-2 protein were associated with clinical

or *in vitro* spread phenotypes. In examining the coding potential of all ten neonatal HSV-2 isolates, we found no protein deletions or truncations encoded by any isolate. These comparisons excluded the viral proteins ICP0, ICP34.5, and ICP4, whose coding sequence is not fully determined in all isolates due to sequencing gaps and/or incomplete assembly at G+C-rich tandem repeats in these regions. Issues in sequencing and assembly of these three proteins have been observed in all prior HTSeq studies of HSV-2 genomes (50, 51, 19). For all other HSV-2 proteins, we found a similar level of coding diversity in this set of neonatal isolates (**Figure 5**) as in a comparable set of 10 adult HSV-2 isolates.

Minority variants expand the potential coding diversity of neonatal HSV-2 isolates

The amino acid (AA) variations described above exist in the consensus genomes of each isolate. Since each round of viral replication creates a population of genomes, we next assessed whether minor variants existed within the viral population of each neonatal HSV-2 isolate. The significant depth of coverage from deep-sequencing of each isolate allowed us to screen for minority variants at every nucleotide position of each genome. We found minority variants in the viral genome population of all 10 neonatal HSV-2 isolates, albeit to a different degree in each isolate (**Figure 6A).** In total, there were 1,452 minority variants present at a frequency of at least 2% of reads at a given locus (see Methods for details). These minority variants were distributed in all genomic regions. Isolates CNS15 and DISS29 had 10-fold higher levels of minority variants than other neonatal isolates (**Figure 6A**), and these variants were often present at a higher frequency or penetrance of the alternative allele than observed in other neonatal or adult HSV-2 isolates (**Figure 6C**)**.** For minority variants located in coding regions, 47% were synonymous and 51.5% were nonsynonymous (**Figure 6B).** We further examined the specific

proteins impacted by these nonsynonymous minority variants (**Figure 6D**), and found that nearly every HSV-2 protein harbored minority variants in at least one neonatal isolate. Only UL3, UL11, UL35, and UL55 were completely devoid of minority variants. Two of these genes (UL3 and UL11) were also devoid of AA variations at the consensus level (**Figure 5**). Overall, the amount of minority variants in these neonatal isolates was similar to what has previously been seen in adult HSV-2 samples, with the exception of the two isolates (CNS15 and DISS29) that had significantly more diverse viral populations.

Coding variations identified between neonatal HSV-2 isolates are associated with CNS disease

To understand how viral genetic variants relate to clinical disease, we assessed whether any of the identified coding variations segregated with clinical disease features. We found that a number of variants were unique to viral isolates obtained from the CSF of neonates with CNS disease (CNS11 and DISS14). These include AA variants at the level of the consensus genome sequence in the HSV-2 UL8 gene (R221S), the UL53 gene that encodes for glycoprotein K (gK, V323M), the US2 gene (F137L), the US4 gene that encodes for glycoprotein G (gG, R338L, S442P, and E574D), and a pair of variants in the HSV-2 US7 gene that encodes glycoprotein I (gI, R159L and P215S) **(Figure 7).** These gI variants exist individually in other viral isolates obtained from the skin of neonates with CNS disease, but are only found together in those HSV-2 isolates obtained from the CSF **(Supplemental Table 2).** Viral isolates from all infants with CNS disease (including disseminated disease with CNS involvement) shared a unique variant in the HSV-2 UL24 gene (A93V). **(Figure 7)**. Viral isolates from infants with disseminated disease with CNS involvement shared a unique variant in the HSV-2 UL20 gene (P129L) **(Figure 7).** Many of

these coding variations that correlate with neonatal CNS disease phenotypes impact viral proteins known to modulate cell-to-cell spread (52–56) and/or contribute to neurovirulence in mouse models of CNS infection (54, 57–63) **(Figure 7 and Table S2)**, however, their role in human disease has not been described. Additional variants in proteins not known to be associated with neurovirulence were also found to be associated with neonatal CNS disease. These represent potentially novel neurovirulence associated proteins that require further validation.

Discussion

The epidemiology of neonatal HSV infection differs greatly from that observed in adults, with over 30% of infected infants experiencing CNS disease, and another 20% experiencing invasive disseminated disease. There is growing evidence that most herpesviruses contain significant genetic variation, including HSV-1 and HSV-2 (19, 22, 50, 51, 64, 65). The potential contributions of viral genetic variation to clinical disease in neonates therefore warrants exploration. Here, we analyze genetic and phenotypic diversity for neonatal HSV-2. We used ten low-passage HSV-2 isolates for which we have detailed information from prior clinical studies on their clinical manifestations at diagnosis and outcomes at one year post-infection (8, 43, 44). We found that neonatal HSV-2 isolates exhibit diverse growth characteristics in culture, with larger average plaque size and enhanced spread through culture being associated with isolates obtained from neonates with CNS disease. Using comprehensive comparative genomics analyses, we further demonstrated that these neonatal HSV-2 isolates contain extensive genetic diversity both within and between hosts. These data revealed several viral genetic variants that are unique to cases of CNS disease, which impact proteins known to contribute to cell-to-cell

spread and/or neurovirulence in mouse models of CNS disease.

Neonatal isolates resulting in CNS disease that were obtained directly from the CSF (CNS11 and DISS14) formed plaques that were significantly larger than all isolates obtained from the skin, and showed enhanced spread through cultured cells **(Figure 1).** Overall, large-plaque-forming isolates were more likely to be derived from patients with CNS disease **(Figure 1)**. It is not surprising that superior cell-to-cell spread would make a virus more fit to enter and exploit the CNS, given that this process requires transmission from surface skin or mucosa into the nervous system, and subsequently neuron-to-neuron spread throughout the brain. Previous work evaluating viral isolates from different body compartments of an HSV-infected neonate also detected differences in plaque size, and found that the most neurovirulent isolates in a mouse model of CNS disease were those obtained from the CSF or brain parenchyma (20). Since the relationship of murine models to disease progression in neonates is unclear (66), we have focused this initial study on the direct comparison of viral genetic features to clinical outcomes and *in vitro* cell-based measures of viral fitness. Further exploration of these neonatal isolates in an appropriate animal model of neonatal disease may reveal further aspects of phenotypic difference (20).

Genomic comparison of these neonatal isolates revealed a wide range of genetic diversity, including at the level of their distinct phylogeny. At the consensus genome level, which reflects the most dominant allele in each viral population, we found that coding differences between strains were as numerous as between previously described sets of adult HSV-2 isolates **(Figure 5)** (19, 50, 51). However at the level of minority variants (MV), which represent rare alleles that exist within each viral population, we found that several viral isolates harbored ten-fold more

MV than others **(Figure 6).** This could be indicative of mixed viral populations or changes in polymerase fidelity (24, 22). As in congenital HCMV infections (25, 28, 30), these data suggested that patients experiencing neonatal infection with HSV-2 may harbor mixed or diverse viral populations that are distinct from those observed in adult infections. These minor genotypes may be selected or genetically isolated in particular niches (e.g. CSF) at a later point in disease, as seen in VZV skin vesicles (34), or by antiviral drug selection, as was recently demonstrated for adult HSV-2 (18). All of the isolates sequenced in this study were collected at the initial time of diagnosis, prior to acyclovir treatment or the development of the infant's immune response over the first year of infection. It would therefore be compelling to examine the viral genome population from subsequent isolates from these patients, to observe if there is any progression or shift in the frequency of MV, or evidence of particular genotypes in different body niches, or signs of selective sweeps due immune selection over time.

Comparison of these 10-full-length neonatal HSV-2 genomes also revealed coding variations associated with CNS disease. We found that isolates obtained from the CSF (CNS11 and DISS14) shared several unique variations, many of these in HSV-2 proteins known to impact cell-to-cell spread and neurovirulence in mouse models of CNS disease. Glycoprotein K (encoded by UL53) is a transmembrane protein important for cytoplasmic virion envelopment and egress. A gK-null virus was found to form smaller plaques *in vitro* and spread less rapidly between neurons in a microfluidics assay as compared to wild-type virus (52). Glycoprotein I (encoded by US7), is a transmembrane protein also known to be important for cell-to-cell spread, including neuron-to-neuron spread (53, 67, 58, 68). Following ocular challenge with HSV in a mouse model of CNS infection, both mutant gK-null (57) and gI-null virus (58) were markedly

restricted in their ability to spread within the eye and from the eye to the nervous system, suggesting each is important for neurovirulence possibly through modulation of cell-to-cell spread (67, 69). Similarly, HSV containing a mutant form of US2 has been shown to form small plaques in culture (54). Following intracerebral injection of either HSV containing a mutant form of US2 (59) or glycoprotein G (encoded by US4) (54), mice exhibit decreased death compared to injection of WT virus. Two additional variants were found to be associated with CNS disease. One variant in UL24 (A93V) was found to be associated with all isolates causing CNS disease (CNS03, CNS11, CNS12, DISS14, CNS15, CNS17, DISS29) **(Figure 7).** UL24-null mutants form small plaques in culture (55), and result in decreased neuronal infection as compared to wild-type virus following ocular challenge in mice (60–62). A variant in UL20 (P129L) was found to be uniquely associated with both isolates causing CNS disease in the context of disseminated infection (DISS14 and DISS29). UL20 encodes for a transmembrane protein that is required for glycosylation and cell surface expression of gK (70). UL20-null mutants form small plaques in culture (56). Additionally, RNA inhibition of UL20 results in decreased rates of encephalitis following HSV footpad injection in mice (63).

Interestingly, isolates obtained directly from the CSF shared the most unique genetic variations. This may be an artifact of the small number of samples in this group, however, the known relevance of the identified proteins to neurovirulence suggest these variations warrant further exploration. The accumulation of neurovirulence-associated variations in isolates obtained from the CSF, but not all isolates causing CNS disease (including isolates obtained from the skin), may indicate that virus able to enter the CNS is genetically distinct from virus contained in other neonatal body compartments. For example, both viral isolates obtained directly from the CSF

contained the unique combination of two US7 variants (R159L and P215S) at the level of the consensus genome. CNS isolates obtained from the skin often contained only one of these gI variants, at the consensus or minority variant level (see **Figure 7, Table S2**). A comparison of skin vs. CSF isolates from the same infant would reveal whether the sequential acquisition of these genetic variants correlates with viral spread from the skin into the CSF.

This comparison of viral genotype to clinical phenotype revealed associations between neonatal CNS disease and several viral protein variants that may impact neurovirulence through modulation of cell-to-cell spread. Although the sample set in this proof of concept study is small, we observed potential patterns that warrant exploration in a larger dataset. It is important to acknowledge that there is limited availability of samples from neonatal infection, both due to the rarity of these infections, and the fragility and limited body size of the infected infants. These natural circumstances lead to minimal sample collection from infected neonates. The finding that CNS-associated isolates exhibit enhanced spread between cells in culture -- particularly those derived directly from the CSF -- suggests that one or more of these variants could be functionally significant. Coding differences in viral proteins not known to contribute to neurovirulence were also found to be associated with neonatal CNS disease, and represent potential novel determinants of invasive infection. These promising results warrant exploration in a larger study, ideally with isolates from multiple time points and/or body sites from each infected infant. This would enable a better understanding of how overall viral genetic diversity contributes to neuroinvasion.

Methods

Virus source

Viruses analyzed were collected from neonates that were culture- or PCR-confirmed to have HSV-2 infection, enrolled in three published clinical studies (8, 43, 44) by the Collaborative Antiviral Study Group (CASG) at the University of Alabama Birmingham. These HSV isolates originated from patients with SEM, CNS, or disseminated (DISS) disease with CNS involvement (**Table 1**). Samples were collected at the time of diagnosis, prior to initiation of acyclovir therapy, from either the cerebrospinal fluid (CSF) or skin. Viral isolates are associated with de-identified clinical information including age, sex, race, and clinical morbidity scores. Enrollment in the original studies was evenly split between males and females, and included both black and white patients. Clinical morbidity score was determined by the CASG after 12 months of life as (1) normal; (2) mild impairment which includes ocular sequelae (keratoconjunctivitis), speech delay, or mild motor delay; (3) moderate impairment which includes hemiparesis, persistent seizure disorder, or developmental delay of less than or equal to 3 months adjusted developmental age; (4) severe impairment which includes microcephaly, spastic quadriplegia, chorioretinitis or blindness, or a serious developmental delay of least 3 months according to the Denver Developmental Assessment Scale. Initial collection of samples was approved by the University of Alabama at Birmingham (UAB) Institutional Review Board. Use of these samples in this study was approved by the UAB, Pennsylvania State University (PSU), and Children's Hospital of Philadelphia Institutional Review Board.

Cell culture

Human lung fibroblast MRC-5 cells (ATCC®, CCL-171) were cultured in minimum essential medium Eagle (MEME, Sigma-Aldrich; M5650) supplemented with 10% fetal bovine serum (FBS, Hyclone; SH30071.03), 2mM L-glutamine (Gibco, A2916801) and 1X penicillin-streptomycin (Gibco; 15140122). African green monkey kidney Vero cells (ATCC®, CCL-81) and human epithelial bone osteosarcoma U2OS cells (ATCC®, HTB-96) were cultured in Dulbecco's modified Eagle's medium with high glucose (DMEM, Hyclone; SH30081.02) supplemented with 10% FBS, 2mM L-glutamine, and 1X penicillin-streptomycin. Cell lines were authenticated by ATCC prior to purchase, and were confirmed to be mycoplasma free throughout experiments by periodic testing (LookOut Mycoplasma, Sigma).

Virus culture

Initial viral isolates (passage 1) were obtained from CASG samples stored at the UAB and shipped to PSU. These stocks were expanded over three serial passages by infecting monolayers of human MRC-5 cells at a MOI of 0.01. Each infection was allowed to progress to complete cytopathic effect (CPE) before harvest (between 50-70 hours). Viral stocks were titered on monolayers of either Vero cells or U2OS cells, for 100 or 48 hours respectively, to allow plaques to develop. Plaque formation was facilitated by limiting viral diffusion with a methylcellulose overlay. Plaque size and morphology were monitored carefully and did not change for any viral isolates over the course of virus stock expansion.

Plaque measurements

After appropriate incubation plaques were stained with 0.5% methylene blue and allowed to dry. Serial 4X brightfield images were collected on an EVOS FL Auto Imaging System and stitched by EVOS software to create an image of the entire well (University of Pennsylvania Cell and Developmental Biology Microscopy Core). No processing was performed. The area of 100 plaques was measured for each viral isolate using ImageJ software.

Genome copy number estimation by quantitative PCR for UL27

DNA was extracted using a PureLink genomic DNA mini kit (ThermoFisher Scientific). Viral genome copy number was determined using an established assay based on real-time PCR using primers and dual-fluorescent probe specific to viral glycoprotein B gene (gB; UL27) (71). Samples were assayed alongside a standard curve of HSV-1 strain F nucleocapsid DNA (72), on an Applied Biosystems 7500 Fast Real-Time PCR system or ViiA 7 Real-Time PCR System.

Viral entry assay

Monolayers of Vero cells were cooled to 4C for 30 minutes prior to infection and 100 pfu of each viral isolate was applied to cell monolayers at 4C for 1h to allow virus binding, after which unbound virus was washed from the cells. Cells were then moved to 37C to allow virus entry. At 0, 10, 20, 30, 45, or 60 minutes, a low-pH citrate buffer was applied to infected cells to inactivate virus that had not penetrated the cellular membrane. At each time point and for each virus, parallel infections were performed without the addition of citrate solution. These served as controls to determine the maximum number of plaques formed. Cell monolayers were washed

and overlaid with methylcellulose. Plaques were scored after 100 hours of incubation**.** Viral entry was quantified as the fraction of plaques formed following citrate buffer application, where 100% is the number of plaques formed on a monolayer not treated with citrate buffer (control).

Single-step and multi-step growth curves

Virus diluted in basic growth media containing 2% FBS was applied to near confluent monolayers of Vero cells and allowed to adsorb for 1 hour. Virus was removed and cells were incubated at 37C for the duration of infection. At 2 hpi, media containing 0.1% human serum was added to reduce the contribution of cell free spread of virus. Single-step growth curves were performed at MOI=5 as defined by titering viral stocks on U2OS cells, and monolayers harvested at 2, 6, 12, and 24 hpi. Multi-step growth curves were performed at MOI=0.001 as defined by titering viral stocks on U2OS cells, and monolayers were harvested at 2, 24, 48, and 72 hpi. Every 24 hours, the supernatant was removed and media containing 0.1% human serum reapplied. At the conclusion of each infection, cells were washed two times with PBS and collected by scraping into an equal volume of media.

Immunocytochemistry

Multi-step growth curves were performed as described above in Vero cell monolayers grown on glass coverslips. To terminate each infection, coverslips were washed with PBS and fixed in 4% paraformaldehyde for 15 minutes at room temperature. Cells were permeabilized with 0.5% Triton-X, blocked in 3% BSA, and incubated with polyclonal rabbit antibodies raised against total HSV (Aligent Dako, B0114). After washing, cells were incubated with fluorophore-conjugated anti-Rabbit secondary antibodies (Invitrogen, A-11008) to mark HSV infected cells,

and counterstained with DAPI to mark cell nuclei. Coverslips were initially visualized with a Leica DM6000 wide field microscope (UPenn Cell and Developmental Microscopy Core). The 5X images were collected on a Photometrics HQ2 high resolution monochrome CCD camera, and processed with LAS AF software. The 10X images were collected on a EVOS FL Auto Imaging System (UPenn Cell and Developmental Microscopy Core) and stitched using EVOS software. Exposure and gain were optimized within each experiment for one virus at the 72-hour time point and applied identically to each image within that experiment. Any subsequent image processing (ImageJ) was applied equally to all images in a given experiment.

Immunoblotting

Whole cell lysates were prepared with 1X LDS sample buffer (NuPage) and equal amounts of lysate were separated by SDS-PAGE. Membranes were immunoblotted with polyclonal rabbit antibodies raised against total HSV (Aligent Dako, B0114), ICP8 (gift from David Knipe), and GAPDH (GeneTex, GTX100118). Proteins were visualized with Pierce ECL Western Blotting Substrate (ThermoFisher Scientific) and detected using a G:Box imaging system (Syngene).

Viral DNA isolation and Illumina sequencing

Viral nucleocapsid DNA was prepared by infecting MRC-5 cells at an MOI ≥5 as previously described (73, 74). Viral nucleocapsid gDNA was sheared using a Covaris M220 sonicator/disruptor under the following conditions: 60s duration, peak power 50, 10% duty cycle, at 4°C. Barcoded sequencing libraries were prepared using the Illumina TruSeq low-throughput protocol according to manufacturer's specifications and as previously described (75, 76). The quality of sequencing libraries was evaluated by Qubit (Invitrogen, CA), Bioanalyzer (Agilent),

and qPCR (KAPA Biosystems). Paired-end sequencing (2 × 300bp length) was done on an Illumina MiSeq, according to manufacturer's recommendations (17pM input).

De novo genome assembly

A consensus genome was assembled for each viral isolate using a previously described Viral Genome Assembly (VirGA) bioinformatics workflow (75). VirGA begins by quality-filtering the MiSeq sequence reads and removing sequences that match the host (human) genome. Thereafter VirGA uses a combination of SSAKE *de novo* assemblies run with differing parameters, which are then combined into a single draft genome using Celera and GapFiller (77–79). After quality-checking and iterative improvement of the genome assembly, annotations were transferred from the HSV-2 reference genome (strain HG52; GenBank NC_001798) to each new genome based on sequence homology (80).

Comparative genomics and phylogenetic analysis

The genomes of all 10 neonatal HSV-2 isolates were combined with all published HSV-2 genomes available in GenBank (see **Supplemental Table 1** for full list) and aligned using MAFFT (81). All previously published HSV-2 genomes were derived from infected adults. The genome-wide alignment used a trimmed genome format (lacking the terminal repeats) to avoid giving undue weight to these duplicated sequences. The MAFFT alignment was used to generate a NeighborNet phylogenetic network in SplitsTree with Uncorrected P distances (82, 39, 83). A diverse subset of ten adult HSV-2 isolates was selected for protein-level comparisons with the ten neonatal isolates: H1227 (Finland), H12212 (Finland), KX574906 (Uganda; strain Uganda.2), KX574874 (Peru; strain Peru.8), KX574883 (Zambia; strain Zambia.4), NC_001798

(Scotland; strain HG52), KP334093 (USA; strain CtSF-R), KP334095 (Wisconsin USA; strain 1192), KF781518 (South Africa; strain SD90e), JX112656 (Texan USA; strain 186). GenBank genomes that lacked ORF and protein annotations were excluded from selection (50) for the protein-based comparisons. ClustalW2 (Larkin et al., 2007) was used to construct pairwise nucleotide alignments between whole genomes and pairwise amino acid alignments for each gene and protein (84). Custom Python scripts were used on these alignments to identify nucleotide and AA differences between samples.

Minority variant detection & quantification

Minority variants (MV) were detected using VarScan v2.2.11 (mpileup2snp and mpileup2indel commands) (85), using the following parameters to differentiate true MV from technical artifacts (34): minimum allele frequency ≥ 0.02 (2%); base call quality ≥ 20; read depth ≥ 100; independent reads supporting minor allele ≥5. MV with directional strand bias ≥ 90% were excluded. The genomic location and potential impact of each MV was assessed using SnpEff and SnpSift (86, 87).

Acknowledgements

We thank Penny Jester and the UAB clinical support team for their assistance in retrieving archival records. The UPenn Cell and Developmental Biology Microscopy Core provided imaging assistance, and Elise M. Peaurori provided technical assistance during early stages of this project. We thank members of the Szpara and Weitzman labs for helpful discussions. Research in the Szpara lab was supported by the startup funds from the Pennsylvania State University (MLS) and a grant with the Pennsylvania Department of Health using Tobacco CURE funds (MLS). The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions. Research in the Weitzman lab was supported in part by a grant from the National Institutes of Health (NS082240 to MDW). LNA was supported the Pediatric Scientist Development Program through grant K12-HD000850 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The published NIAID CASG clinical studies of neonatal HSV infection were funded by a contract (N01-AI-62554) with the NIAID Development and Applications Branch and by grants from the General Clinical Research Center Program (RR-032) and the state of Alabama.

Author Contributions

We describe contributions to the paper using the CRediT taxonomy (88). Writing – Original Draft: LNA, CB, ADF, and MLS; Writing – Review & Editing: LNA, MDW, and MLS; Conceptualization: LNA and MLS; Investigation: LNA, CB, ADF, DWR, UP; Resources: DWK,

MNP, and RJW; Funding Acquisition: MLS, RJW, and MDW.

Declaration of Interests

The authors declare no competing interests.

Tables

Table 1: Clinical characteristics associated with HSV-2 isolates from ten patients.

Clinical Isolate*	Clinical Disease at Diagnosis	Sample Source	Morbidity Score - Mental	Morbidity Score - Motor	Patient Age at Disease Onset (days)	Gestational Age at Time of Birth (weeks)	Patient Sex & Race
CNS11	CNS	CSF	4	4	12	37	M, W
DISS14	DISS + CNS	CSF	2	2	7	39	M, W
CNS03	CNS	SKIN	4	4	17	37	M, W
CNS15	CNS	SKIN	3	4	19	36	M, W
CNS17	CNS	SKIN	4	4	17	40	M, W
DISS29	DISS + CNS	SKIN	3	3	5	38	F, B
CNS12	CNS	SKIN	4	4	16	41	F, W
SEM02	SEM	SKIN	1	1	5	38	F, W
SEM13	SEM	SKIN	4	4	11	27	F, B
SEM18	SEM	SKIN	2	2	17	37	F, W

* Clinical isolate order based on Figure 1.

Table 2. Genome sequencing statistics for neonatal HSV-2 strains.

Sample Name	Disease Type	Raw sequence reads	Used for assembly	Complete ORFs	Depth > 100
RW02	SEM	4,767,684	3,679,189	69	95%
RW03	CNS	3,947,550	3,072,706	68	96%
RW11	CNS	3,503,282	2,757,441	70	96%
RW12	CNS	3,875,582	2,823,012	71	96%
RW13	SEM	3,393,466	2,437,202	62	89%
RW14	DISS	3,846,859	2,951,334	71	97%
RW15	CNS	6,074,071	4,425,705	69	97%
RW17	CNS	3,105,207	2,274,958	70	96%
RW18	SEM	16,895,584	12,236,021	71	98%
RW29	DISS	1,415,376		72	99%

Figures

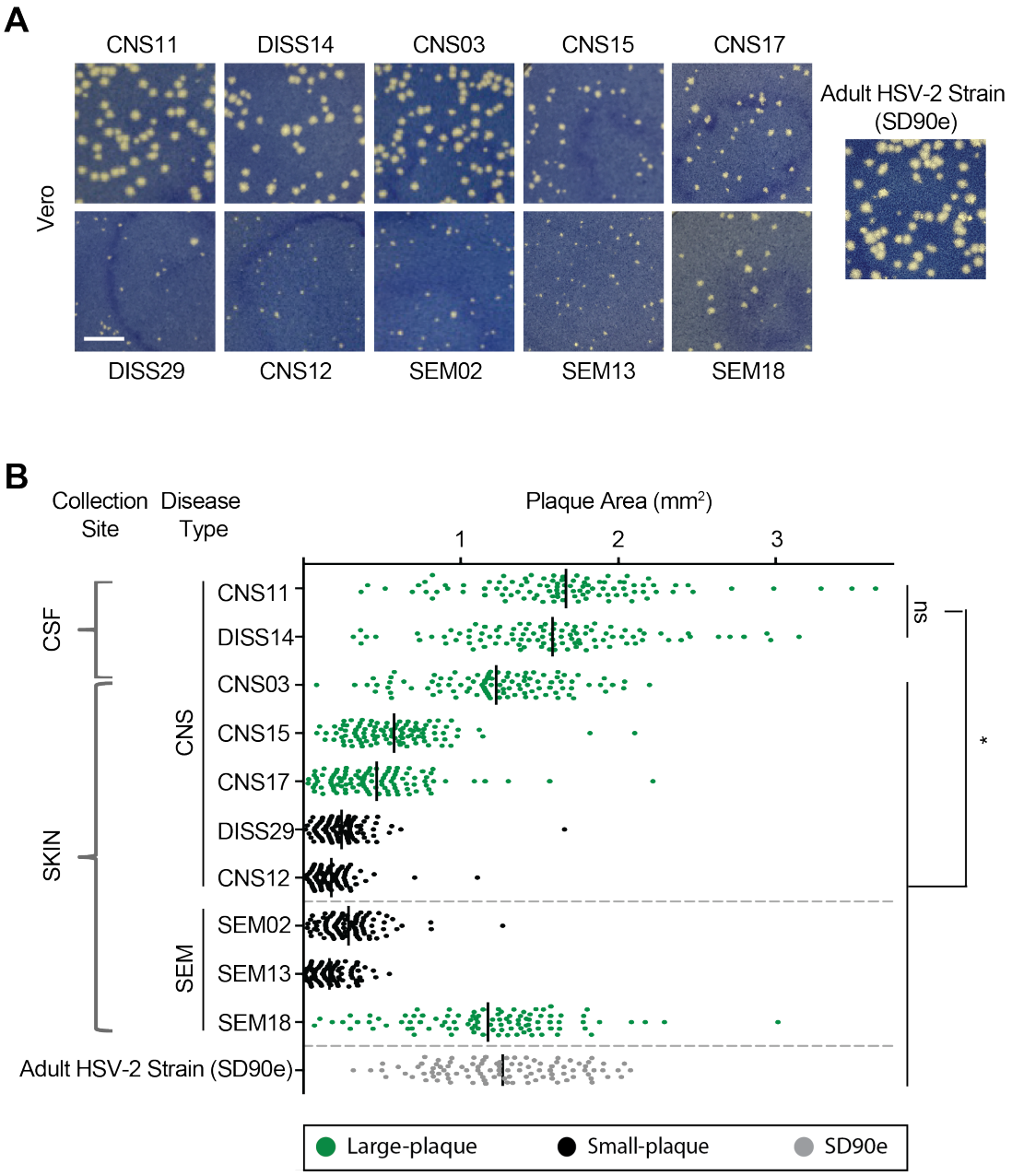


Figure 1. Neonatal HSV-2 isolates generate different sized plaques in culture.

(**A**) Representative plaques are shown after neonatal viruses were allowed to incubate for 100 hours on Vero cells. Plaques formed by the low-passage adult HSV-2 strain SD90e are shown for

comparison. Cell monolayers are stained with methylene blue. Scale bar = 5mm. (**B**) Quantification of plaque area on Vero cells for each isolate. Dots represent 100 individually measured plaques and black bar represents the mean. Largest average plaque areas are indicated in green, and smallest plaque areas indicated in black. SD90e is indicated in gray. Clinical disease is indicated as CNS (inclusive of CNS only and DISS + CNS disease) or SEM, and site of virus collection indicated as CSF or SKIN. Each green isolate is statistically larger than each black isolate. Black isolates are not statistically different from one another. Additionally, each CSF-derived isolate is statistically larger than all other isolates shown. For all statistics, $p<0.05$ by one-way ANOVA followed by Holm-Sidak's multiple comparisons test.

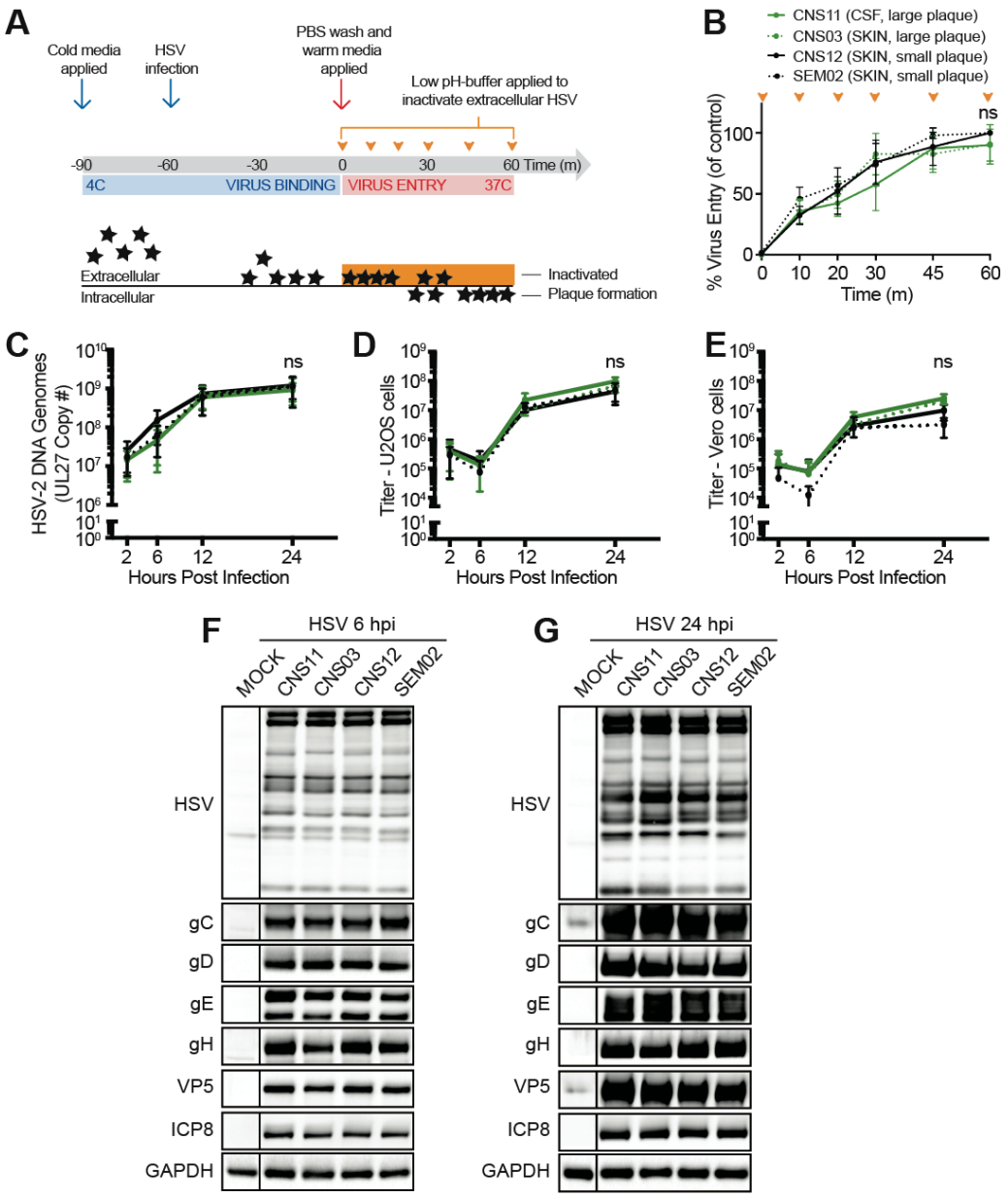


Figure 2. Increased plaque size in culture is not influenced by viral entry or viral DNA replication.

Viral growth characteristics were compared for several representative neonatal HSV-2 isolates, including large-plaque-formers (green) and small-plaque-formers (black). **(A-B)** Viral entry kinetics. The rate of virus penetration into Vero cells was compared. **(A)** Cells were cooled to 4C

for 30 minutes prior to infection. Each viral isolate was applied to cell monolayers at 4C for 1h to allow virus binding, then moved to 37C to allow virus entry. Extracellular virus was inactivated by the application of a low-pH citrate buffer at the times indicated (orange arrowheads). Cell monolayers were washed and overlaid with methylcellulose. Plaques were scored after 100 hours of incubation. **(B)** Viral entry is quantified as the fraction of plaques formed following citrate buffer application, where 100% is the number of plaques formed on a monolayer not treated with citrate buffer (control). These data represent three independent experiments. Two-way ANOVA followed by Tukey's multiple comparison test was applied. **(C-E)** Viral DNA replication kinetics. The rate of viral DNA replication in Vero cells was compared. Confluent Vero cell monolayers were infected at MOI = 5 and incubated in the presence of 0.1% human serum for the time points indicated. The cell monolayer was harvested at each time point. **(C)** The quantity of viral genomes present was evaluated by qPCR for UL27. Viral titer was evaluated by plaque formation on U2OS **(D)** or Vero **(E)** cells. These data represent three independent experiments. Two-way ANOVA followed by Tukey's multiple comparison test was applied. Confluent Vero cell monolayers were infected at MOI=5 for 6h **(F)** or 24 **(G)** hours. Whole cell lysates were subjected to immunoblot analysis with HSV antibodies**.**

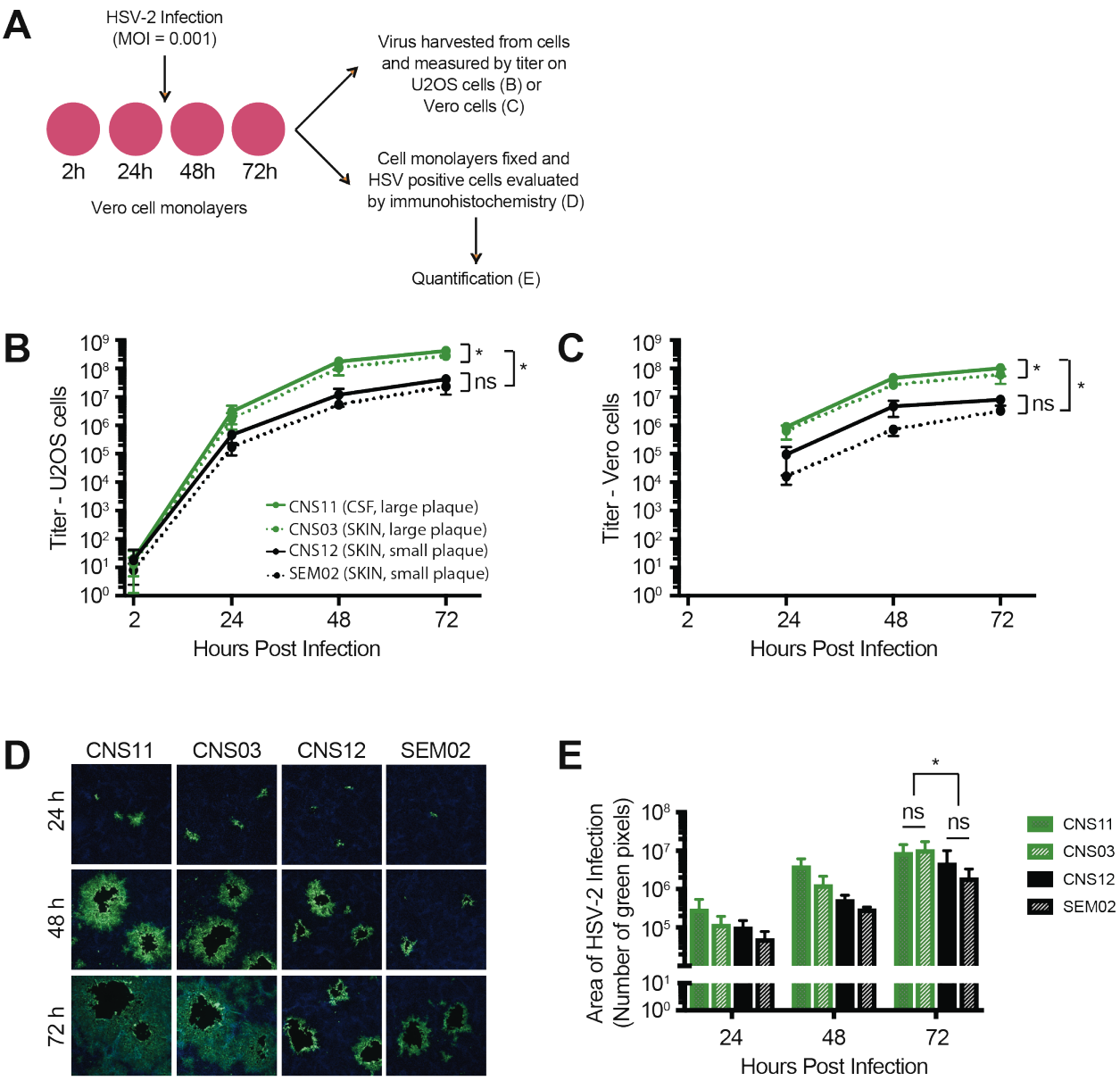


Figure 3. Enhanced viral cell-to-cell spread contributes to increased plaque size in culture.

(**A**) The rate of viral spread in Vero cells was compared for several representative neonatal isolates, including large-plaque-formers (green) and small-plaque-formers (black). Confluent Vero cell monolayers were infected at MOI = 0.001 for the time points indicated, in the presence of 0.1% human serum replenished every 24 hours. (**B-C**) The Vero cell monolayer was harvested

at each time point and viral titer was evaluated by plaque formation on U2OS cells **(B)** or Vero cells **(C).** These data represent three independent experiments. Two-way ANOVA followed by Tukey's multiple comparison test, *$p<0.0001$ at 72h. **(D)** In parallel experiments, the Vero cell monolayer was fixed with 4% PFA, and HSV positive cells (green) were evaluated at each time point by immunofluorescence. Cell nuclei are counterstained with DAPI (blue). All images 5X. Exposure, gain, and other image capture settings were optimized for the 72h time point of CNS11 and then applied uniformly for each image. Images are representative of three independent experiments. Images of the entire 10 mm coverslips were then captured and stitched to create a composite image (see **Supplemental Figure 2**). (**E**) The total number of green pixels was quantified for the entire image. Two-way ANOVA followed by Tukey's multiple comparison test, *$p<0.05$ at 72h.

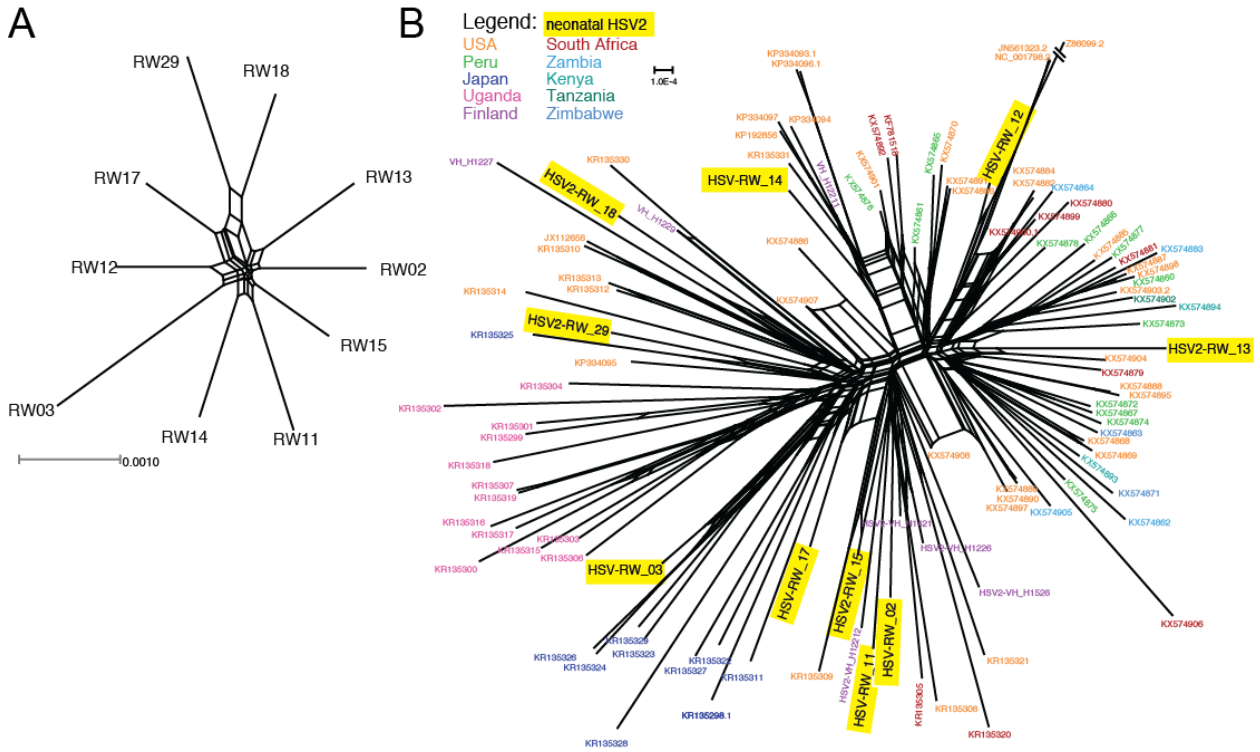


Figure 5. Genetic relatedness of neonatal HSV2 strains

A phylogenetic network among neonatal HSV-2 genomes (**A**), or neonatal and all available adult HSV-2 genomes (**B**), was constructed using SplitsTree4. Country of origin is indicated for the adult isolates (**B**), to demonstrate the inter-mingling of viral isolates relative to their geographic origin. The neonatal HSV-2 genomes appear to be genetically distinct from one another (**A**) and to encompass a broad range of known HSV-2 genetic diversity.

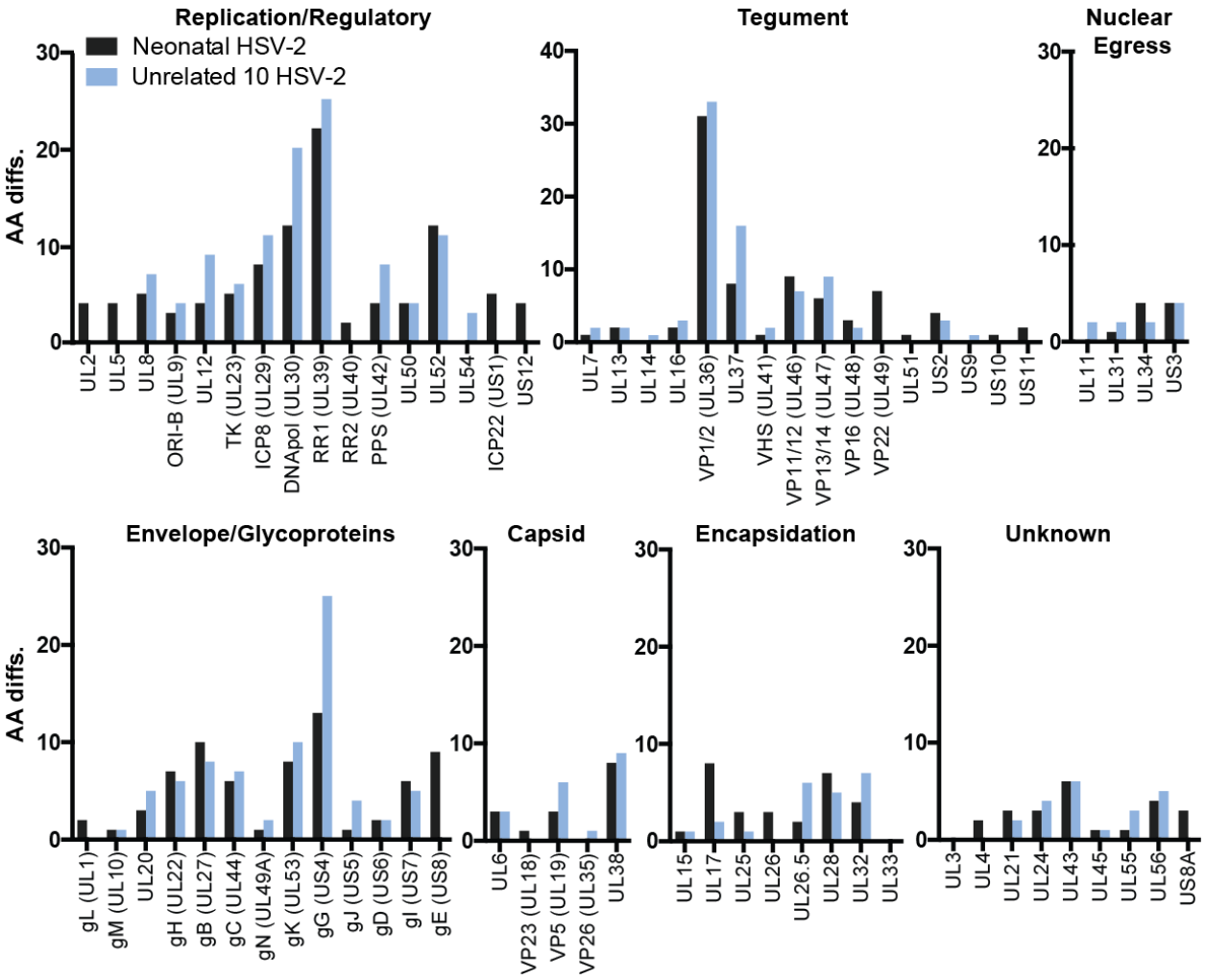


Figure 6: Amino acid differences between consensus genomes of neonatal HSV-2 isolates.

HSV-2 proteins are grouped by function, and the number of AA differecnes between 10 neonatal isolates (black) are compared to those of 10 adult HSV-2 isolates (blue). The viral proteins with AA differences among neonatal but not the adult isolates shown here (UDG (UL2), UL5, ICP22 (US1), ICP47 (US12), US10, US11, gL (UL1), UL26, and UL4) do contain AA variations in other published HSV-2 strains, indicating that these are not universally invariant in adult isolates (50, 51, 19).

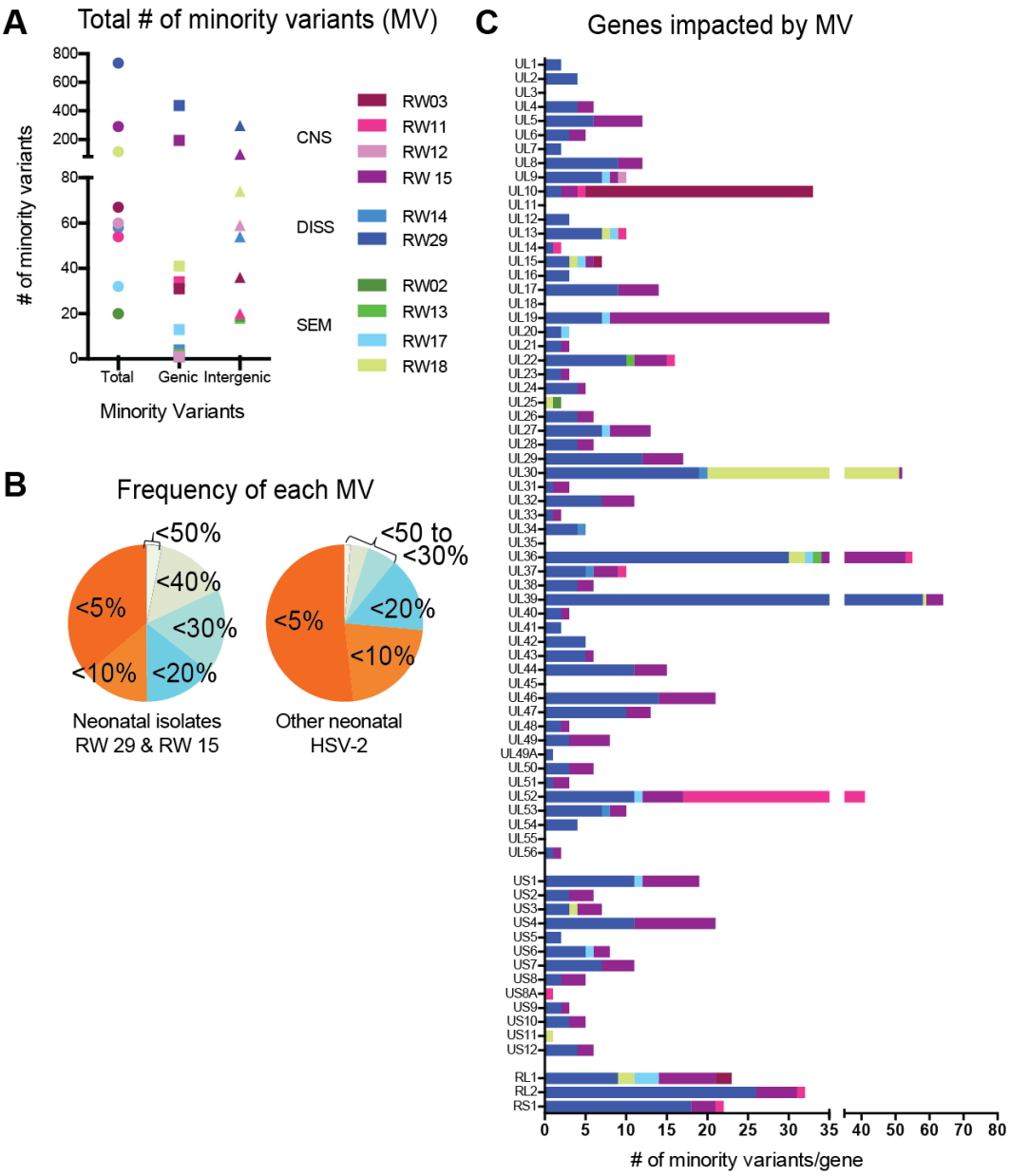


Figure 7: Minority variants present in neonatal HSV-2 genome populations.

(**A**) Plot indicates the total number of minority variants (MV) observed in each neonatal isolate. DISS29 and CNS15 have 10-fold more minority variants than other neonatal strains, which is particularly noticeable for those MV that are located in coding, or genic, sequences. (**B**) Pie charts show the overall frequency or peentrance of each minority variant. DISS29 and CNS15

have many MV that exist at a high frequency or penetrance, while the penetrance of MV alleles in most other neonatal isolates is low. (**C**) Stacked histograms show the number of nonsynonymous MV located in each HSV2 coding sequence (gene). Color coding of stacked histogram bars is the same as shown in (**A**).

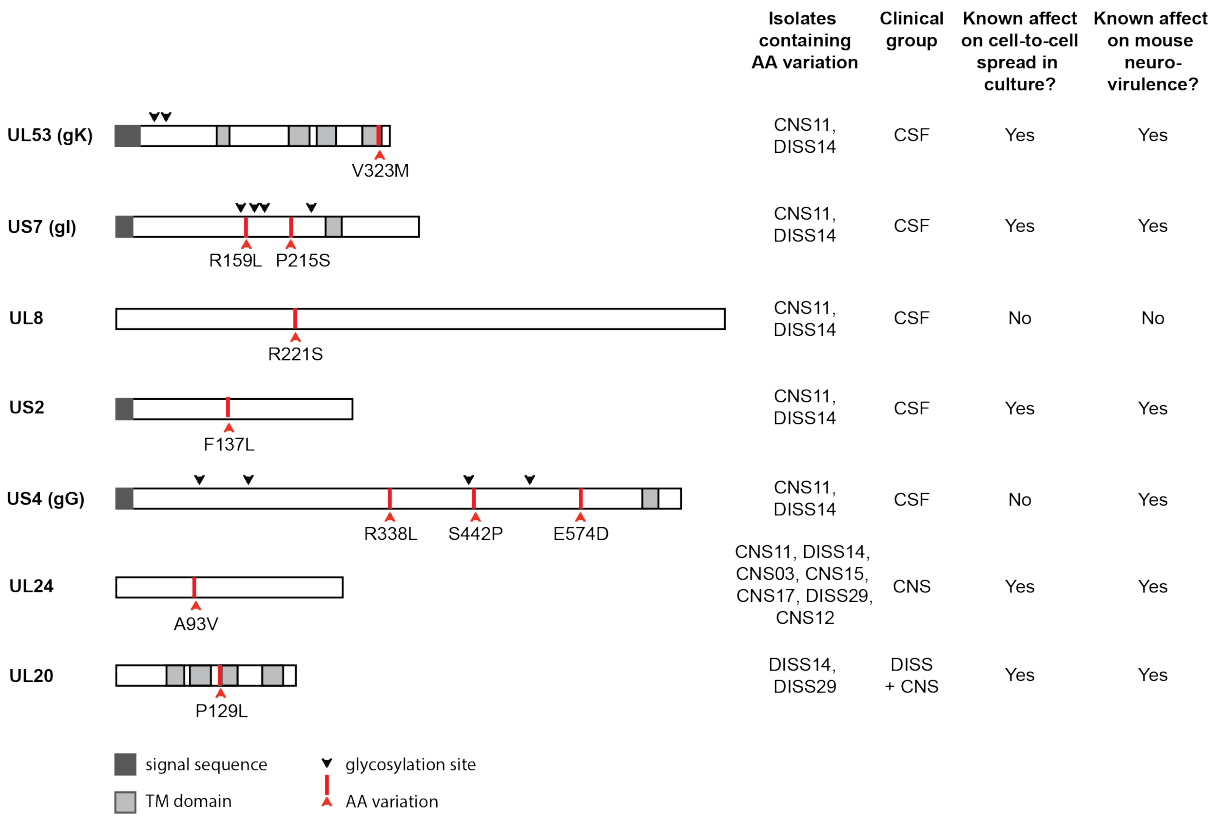
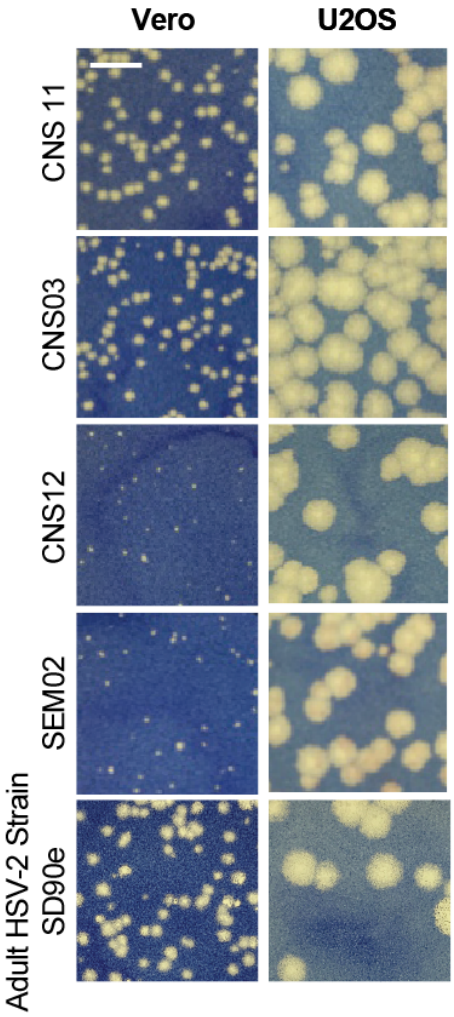


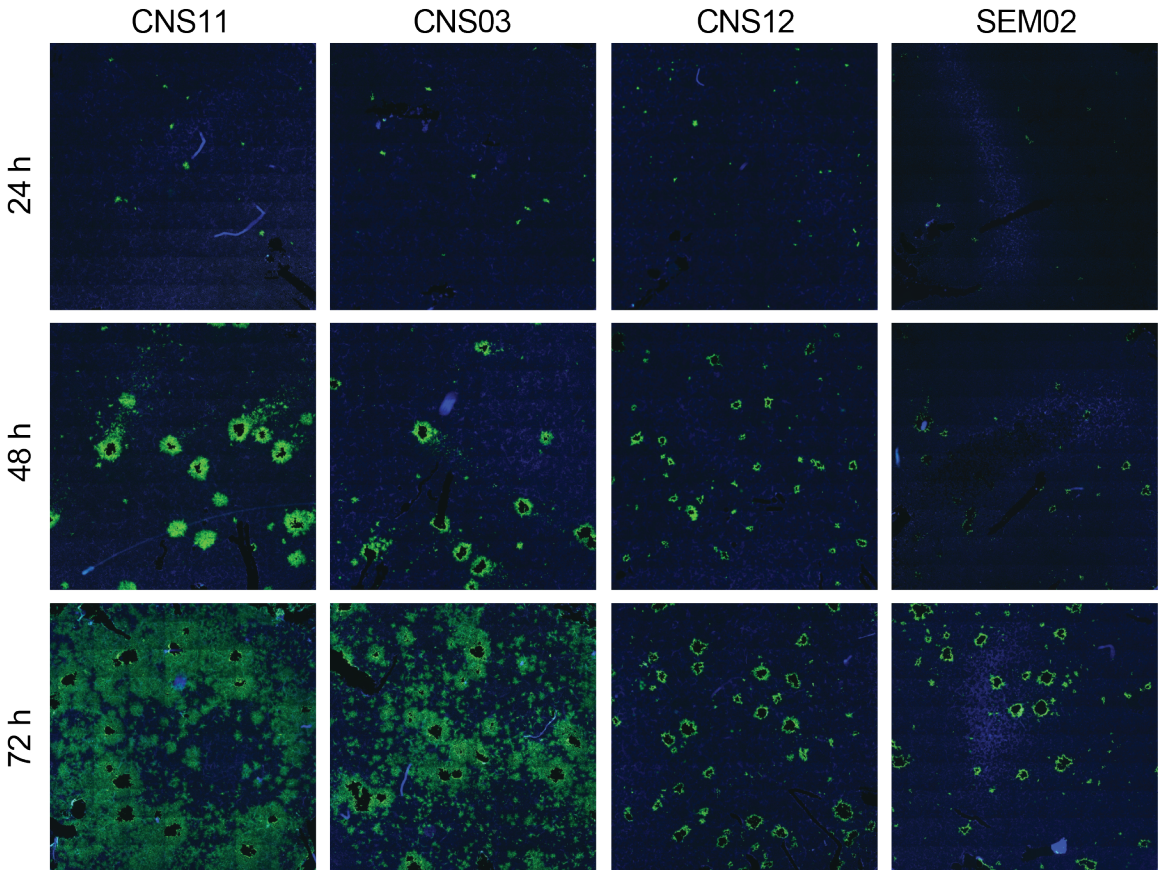
Figure 8. Genetic variations associated with neonatal CNS disease identified in proteins known to contribute to cell-to-cell spread and neurovirulence. The domain structure of each HSV protein is based on published literature for both HSV-1 and -2. More detailed information regarding cell-to-cell spread and neurovirulence can be found in **Supplemental Table 2**.

Supplementary Figures



Supplemental Figure 1. U2OS cells support large plaque formation by all isolates.

Representative plaques are shown after neonatal viruses were allowed to incubate for 100h on Vero **(A)** or USOS cells **(B)**. All neonatal isolates were capable of forming large plaques on U2OS cells, which lack innate sensing of viral infection through the STING pathway (48).



Supplemental Figure 2. Cell-to-cell spread differences between neonatal HSV-2 isolates.

Confluent Vero cell monolayers were infected at MOI = 0.001 for the time points indicated, in the presence of 0.1% human serum. HSV positive cells (green) were evaluated at each time point by immunofluorescence. Cell nuclei are counterstained with DAPI (blue). Serial 10X images were obtained on an EVOS FL Auto Imaging System and stitched together to create an image of the entire experimental coverslip. These images were quantified in **Figure 3E**.

Supplemental Table 1

Virus Isolate	Country (with location detail, i available)	GenBank Accession #	References
H1227*	Finland	In process	In process
H12212*	Finland	In process	In process
Uganda.2*	Uganda	KX574906	(19)
Peru.8*	Peru	KX574874	(19)
Zambia.4*	Zambia	KX574883	(19)
HG52*	Scotland	NC_001798	(89)
CtSF-R*	USA	KP334093	(51)
1192*	Wisconsin, USA	KP334095	(51)
SD90e*	Carletonville, South Africa	KF781518	(47, 90)
186*	Texas USA	JX112656	(91)
Peru.1	Peru	KX574860	(19, 92)
Peru.12	Peru	KX574861	(19, 92)
Zambia.3	Zambia	KX574862	(19, 92)
Zimbabwe.2	Zimbabwe	KX574863	(19, 92)
Zambia.2	Zambia	KX574864	(19, 92)
Peru.9	Peru	KX574865	(19, 92)
Peru.11	Peru	KX574866	(19, 92)
Peru.10	Peru	KX574867	(19, 92)
USA.8a	USA	KX574868	(19, 92)
USA.8b	USA	KX574869	(19, 92)
USA.1b	USA	KX574870	(19, 92)
Zimbabwe.1	Zimbabwe	KX574871	(19, 92)
Peru.7	Peru	KX574872	(19, 92)
Peru.2	Peru	KX574873	(19, 92)
Peru.8	Peru	KX574874	(19, 92)
Peru.6	Peru	KX574875	(19, 92)
Peru.4	Peru	KX574876	(19, 92)
Peru.3	Peru	KX574877	(19, 92)
Peru.5	Peru	KX574878	(19, 92)
South Africa.2	South Africa	KX574879	(19, 92)
South Africa.3	South Africa	KX574880	(19, 92)
South Africa.4	South Africa	KX574881	(19, 92)
USA.3a	USA	KX574882	(19, 92)

Virus Isolate	**Country (with location detail, i available)**	**GenBank Accession #**	**References**
Zambia.4	Zambia	KX574883	(19, 92)
USA.3b	USA	KX574884	(19, 92)
USA.2b	USA	KX574885	(19, 92)
USA.10	USA	KX574886	(19, 92)
USA.4a	USA	KX574887	(19, 92)
USA.7a	USA	KX574888	(19, 92)
USA.6a	USA	KX574889	(19, 92)
USA.6b	USA	KX574890	(19, 92)
USA.5a	USA	KX574891	(19, 92)
South Africa.1	South Africa	KX574892	(19, 92)
Kenya.2	Kenya	KX574893	(19, 92)
Kenya.1	Kenya	KX574894	(19, 92)
USA.7b	USA	KX574895	(19, 92)
USA.5b	USA	KX574896	(19, 92)
USA.11	USA	KX574897	(19, 92)
USA.4b	USA	KX574898	(19, 92)
Uganda.1	Uganda	KX574899	(19, 92)
2011-34957	Uganda	KX574900	(19, 92)
USA.9	USA	KX574901	(19, 92)
Tanzania.1	Tanzania	KX574902	(19, 92)
USA.1a	USA	KX574903	(19, 92)
USA.2a	USA	KX574904	(19, 92)
Zambia.1	Zambia	KX574905	(19, 92)
Uganda.2	Uganda	KX574906	(19, 92)
2006-38659	USA	KX574907	(19, 92)
2011-21400	USA	KX574908	(19, 92)
HG52	UK	NC_001798	
HG52	UK	JN561323	
CtSF	USA	KP334097	(51)
COH 3818	USA	KP334096	(51)
1192	Wisconsin, USA	KP334095	(51)
GSC-56	USA	KP334094	(51)
CtSF-R	USA	KP334093	(51)
333	Texas, USA	KP192856	(51)
SD90e	South Africa	KF781518	(47)
186		JX112656	
HG52	UK	Z86099	(93)
BethesdaP5	Maryland, USA	KR135330	(50)

Virus Isolate	**Country (with location detail, i available)**	**GenBank Accession #**	**References**
333-R519	USA	KR135331	(50)
JA9	Japan	KR135329	(50)
JA8	Japan	KR135328	(50)
JA7	Japan	KR135327	(50)
JA6	Japan	KR135326	(50)
JA5	Japan	KR135325	(50)
JA3	Japan	KR135324	(50)
JA2	Japan	KR135323	(50)
JA1	Japan	KR135322	(50)
89-390	Massachusetts, USA	KR135321	(50)
SD66	South Africa	KR135320	(50)
UG/BID-G19091/D39765	Rakai, Uganda	KR135319	(50)
UG/BID-G19090/D39650	Rakai, Uganda	KR135318	(50)
UG/BID-G19089/A76832	Rakai, Uganda	KR135317	(50)
UG/BID-G19088/G75809	Rakai, Uganda	KR135316	(50)
UG/BID-G19087/J32715	Rakai, Uganda	KR135315	(50)
US/BID-G19086/7444-1996-25809	Washington, USA	KR135314	(50)
US/BID-G19085/9335-2007-14	Washington, USA	KR135313	(50)
US/BID-G19084/9335-2005-576	Washington, USA	KR135312	(50)
US/BID-G19083/10883-2001-13347	Washington, USA	KR135311	(50)
US/BID-G19082/44-319857	Maryland, USA	KR135310	(50)
US/BID-G19081/44-419851	Maryland, USA	KR135309	(50)
US/BID-G19080/44-419851	Maryland, USA	KR135308	(50)
UG/BID-G19079/J09622	Rakai, Uganda	KR135307	(50)
UG/BID-G19078/A76191	Rakai, Uganda	KR135306	(50)
UG/BID-G19077/K39924	Rakai, Uganda	KR135305	(50)
UG/BID-G19076/H00066	Rakai, Uganda	KR135304	(50)
UG/BID-G44423/L22861	Rakai, Uganda	KR135303	(50)
UG/BID-G19074/M1119	Rakai, Uganda	KR135302	(50)
UG/BID-G19073/F70764	Rakai, Uganda	KR135301	(50)

Virus Isolate	**Country (with location detail, i available)**	**GenBank Accession #**	**References**
UG/BID-G19072/D30613	Rakai, Uganda	KR135300	(50)
UG/BID-G19071/M22987	Rakai, Uganda	KR135299	(50)
US/BID-G19070/8937-1-3336	Washington, USA	KR135298	(50)
VH_H12211	Finland	KY922725	In process
VH_H12212	Finland	KY922726	In process
VH_H1226	Finland	KY922720	In process
VH_H1227	Finland	KY922721	In process
VH_H1229	Finland	KY922722	In process
VH_H1421	Finland	KY922723	In process
VH_H1526	Finland	KY922724	In process

Supplemental Table 2

HSV-2 Gene	Genetic Variation	Isolates Containing Variation	Clinical Group	Predicted effect of variation on protein (SMART Analysis)	Protein Function	Neurovirulence
UL53 (gK)	V323M	CNS11, DISS14	All CSF Isolates	Loss of one TM domain	Glycoprotein at virion and cell surface; interacts with gB and UL20 (94); important for cytoplasmic envelopment, viral egress, and virus-induced cell fusion	gK null virus exhibits poor neuronal spread in culture including decreased retrograde and anterograde transport (52); reduced spread within eye and to nervous system, as well as increased survival, following ocular challenge in mice (69)
US7 (gI)	R159L + P215S	Both: CNS11, DISS14 R159L only: CNS03, SEM13, CNS15*, DISS29* P215S only: CNS15*, DISS29	All CSF Isolates	R159L adds one TM domain; P215S results in no identified change	Glycoprotein at virion and cell surface; forms heterodimer with gE to form viral Fc receptor (95); required for cell-to-cell spread of virus in epithelial and neuronal cells (53, 67);	gI null virus exhibits poor neuronal spread in culture; reduced spread in retina and to retinorecipient areas of the brain following ocular challenge in mice (67)
UL8	R221S	CNS11, DISS14	All CSF Isolates	No predicted change	Putative primase subunit of helicase primase complex; required for unwinding of viral DNA	None identified
US2	F137L	CNS11, DISS14	All CSF Isolates	No predicted change	Tegument protein; binds and phosphorylates TAK1 to positively regulate NFkB signaling (96)	Virus containing mutated US2 forms small plaques in culture (54); Mutant virus containing a deletion of both US2 and PK result in decreased death following intracerebral injection in mice (59)
US4 (gG)	R338L, S442P, E574D	CNS11, DISS14	All CSF Isolates	No predicted change	Binds and enhances function of chemokines (97); enhances neurotrophin-dependent axonal growth in culture (98)	Virus containing mutated gG results in decreased death following intracerebral injection in mice (54)
UL24	A93V	CNS03, CNS11, CNS12, DISS14, CNS15, CNS17, DISS29	All CNS Isolates	No predicted change	Membrane-associated nuclear protein; associated with dispersal of nucleolin during infection (99)	UL24 null virus forms small plaques in culture (55); decreased viral load in eye and nervous system following ocular challenge in mice (60–62)
UL20	P129L	DISS14, DISS29	All DISS Isolates w/ CNS Involved	Change in position of three TM domains	Membrane-associated protein interacts with gK and gB (94); required for gK glycosylation, cell surface expression (70), and virus-induced cell fusion (56); important for viral egress (100)	UL20 null virus forms small plaques in culture (56); RNA inhibition of UL20 results in decreased rates of encephalitis following HSV footpad infection in mice (63)

* present as minority variant

References

1. 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.

2. Stephenson-Famy A, Gardella C. 2014. Herpes Simplex Virus Infection During Pregnancy. Obstet Gynecol Clin North Am 41:601–614.

3. 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. Jama 289:203–209.

4. Corey L, Wald A. 2009. Maternal and Neonatal HSV Infections. N Engl J Med 361:1376–1385.

5. Bradley H, Markowitz LE, Gibson T, McQuillan GM. 2014. Seroprevalence of herpes simplex virus types 1 and 2--United States, 1999-2010. J Infect Dis 209:325–333.

6. James SH, Kimberlin DW. 2014. Neonatal Herpes Simplex Virus Infection. Clin Perinatol.

7. Pinninti SG, Kimberlin DW. 2014. Preventing Herpes Simplex Virus in the Newborn. Clin Perinatol 41:945–955.

8. Kimberlin DW, Lin CY, Jacobs RF, Powell DA, Frenkel LM, Gruber WC, Rathore M, Bradley JS, Diaz PS, Kumar M, Arvin AM, Gutierrez K, Shelton M, Weiner LB, Sleasman JW, de Sierra TM, Soong SJ, Kiell J, Lakeman FD, Whitley RJ, National Institute of Allergy and Infectious Diseases Collaborative Antiviral Study Group. 2001. Natural history of neonatal herpes simplex virus infections in the acyclovir era. Pediatrics 108:223–229.

9. Kimberlin DW, Whitley RJ, Wan W, Powell DA, Storch G, Ahmed A, Palmer A, Sánchez PJ, Jacobs RF, Bradley JS. 2011. Oral acyclovir suppression and neurodevelopment after neonatal herpes. N Engl J Med 365:1284–1292.

10. Casanova J-L. 2015. Human genetic basis of interindividual variability in the course of infection. Proc Natl Acad Sci 201521644.

11. Zhang S-Y, Jouanguy E, Ugolini S, Smahi A, Elain G, Romero P, Segal D, Sancho-Shimizu V, Lorenzo L, Puel A, Picard C, Chapgier A, Plancoulaine S, Titeux M, Cognet C, von Bernuth H, Ku C-L, Casrouge A, Zhang X-X, Barreiro L, Leonard J, Hamilton C, Lebon P, Heron B, Vallee L, Quintana-Murci L, Hovnanian A, Rozenberg F, Vivier E, Geissmann F, Tardieu M, Abel L, Casanova J-L. 2007. TLR3 deficiency in patients with herpes simplex encephalitis. Science 317:1522–1527.

12. Fields BN, Byers K. 1983. The Genetic Basis of Viral Virulence. Philos Trans R Soc Lond B

Biol Sci 303:209–218.

13. Gonzalez-Scarano F, Beaty B, Sundin D, Janssen R, Endres MJ, Nathanson N. 1988. Genetic determinants of the virulence and infectivity of La Crosse virus. Microb Pathog 4:1–7.

14. Tscherne DM, García-Sastre A. 2011. Virulence determinants of pandemic influenza viruses. J Clin Invest 121:6–13.

15. Müller V, Fraser C, Herbeck JT. 2011. A Strong Case for Viral Genetic Factors in HIV Virulence. Viruses 3:204–216.

16. Robinson CM, Jesudhasan PR, Pfeiffer JK. 2014. Bacterial Lipopolysaccharide Binding Enhances Virion Stability and Promotes Environmental Fitness of an Enteric Virus. Cell Host Microbe 15:36–46.

17. Yoshii K, Sunden Y, Yokozawa K, Igarashi M, Kariwa H, Holbrook MR, Takashima I. 2014. A Critical Determinant of Neurological Disease Associated with Highly Pathogenic Tick-Borne Flavivirus in Mice. J Virol 88:5406–5420.

18. Minaya MA, Jensen T, Goll J, Korom M, Datla SH, Belshe RB, Morrison LA. 2017. Molecular evolution of herpes simplex virus 2 complete genomes: Comparison between primary and recurrent infections. J Virol JVI.00942-17.

19. Johnston C, Magaret A, Roychoudhury P, Greninger AL, Cheng A, Diem K, Fitzgibbon MP, Huang M, Selke S, Lingappa JR, Celum C, Jerome KR, Wald A, Koelle DM. 2017. Highly conserved intragenic HSV-2 sequences: Results from next-generation sequencing of HSV-2 U L and U S regions from genital swabs collected from 3 continents. Virology 510:90–98.

20. Bower JR, Mao H, Durishin C, Rozenbom E, Detwiler M, Rempinski D, Karban TL, Rosenthal KS. 1999. Intrastrain variants of herpes simplex virus type 1 isolated from a neonate with fatal disseminated infection differ in the ICP34.5 gene, glycoprotein processing, and neuroinvasiveness. J Virol 73:3843–53.

21. Mao H, Rosenthal KS. 2003. Strain-Dependent Structural Variants of Herpes Simplex Virus Type 1 ICP34.5 Determine Viral Plaque Size, Efficiency of Glycoprotein Processing, and Viral Release and Neuroinvasive Disease Potential. J Virol 77:3409–3417.

22. Renner DW, Szpara ML. 2017. The Impacts of Genome-wide Analyses on our Understanding of Human Herpesvirus Diversity and Evolution. J Virol.

23. Houldcroft CJ, Beale MA, Breuer J. 2017. Clinical and biological insights from viral genome sequencing. Nat Rev Microbiol 15:183–192.

24. Houldcroft CJ, Bryant JM, Depledge DP, Margetts BK, Simmonds J, Nicolaou S, Tutill HJ, Williams R, Worth AJJ, Marks SD, Veys P, Whittaker E, Breuer J. 2016. Detection of Low Frequency Multi-Drug Resistance and Novel Putative Maribavir Resistance in Immunocompromised Pediatric Patients with Cytomegalovirus. Front Microbiol 7.

25. Renzette N, Bhattacharjee B, Jensen JD, Gibson L, Kowalik TF. 2011. Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants. PLoS Pathog 7:1–14.

26. 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.

27. Renzette N, Pokalyuk C, Gibson L, Bhattacharjee B, Schleiss MR, Hamprecht K, Yamamoto AY, Mussi-Pinhata MM, Britt WJ, Jensen JD, Kowalik TF. 2015. Limits and patterns of cytomegalovirus genomic diversity in humans. Proc Natl Acad Sci 201501880.

28. Renzette N, Kowalik TF, Jensen JD. 2016. On the relative roles of background selection and genetic hitchhiking in shaping human cytomegalovirus genetic diversity. Mol Ecol 25:403–413.

29. Renzette N, Pfeifer SP, Matuszewski S, Kowalik TF, Jensen JD. 2017. On the Analysis of Intrahost and Interhost Viral Populations: Human Cytomegalovirus as a Case Study of Pitfalls and Expectations. J Virol 91:e01976-16.

30. Pokalyuk C, Renzette N, Irwin KK, Pfeifer SP, Gibson L, Britt WJ, Yamamoto AY, Mussi-Pinhata MM, Kowalik TF, Jensen JD. 2017. Characterizing human cytomegalovirus reinfection in congenitally infected infants: an evolutionary perspective. Mol Ecol.

31. Sijmons S, Van Ranst M, Maes P. 2014. Genomic and Functional Characteristics of Human Cytomegalovirus Revealed by Next-Generation Sequencing. Viruses 6:1049–1072.

32. Sijmons S, Thys K, Mbong Ngwese M, Van Damme E, Dvorak J, Van Loock M, Li G, Tachezy R, Busson L, Aerssens J, Van Ranst M, Maes P. 2015. High-Throughput Analysis of Human Cytomegalovirus Genome Diversity Highlights the Widespread Occurrence of Gene-Disrupting Mutations and Pervasive Recombination. J Virol 89:7673–7695.

33. Lassalle F, Depledge DP, Reeves MB, Brown AC, Christiansen MT, Tutill HJ, Williams RJ, Einer-Jensen K, Holdstock J, Atkinson C, Brown JR, van Loenen FB, Clark DA, Griffiths PD, Verjans GMGM, Schutten M, Milne RSB, Balloux F, Breuer J. 2016. Islands of linkage in an ocean of pervasive recombination reveals two-speed evolution of human cytomegalovirus genomes. Virus Evol 2:vew017.

34. Depledge DP, Kundu S, Jensen NJ, Gray ER, Jones M, Steinberg S, Gershon A, Kinchington PR, Schmid DS, Balloux F, Nichols RA, Breuer J. 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.

35. Sakaoka H, Kurita K, Iida Y, Takada S, Umene K, Kim YT, Ren CS, Nahmias AJ. 1994. 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.

36. Brown J. 2004. Effect of gene location on the evolutionary rate of amino acid substitutions in herpes simplex virus proteins. Virology 330:209–20.

37. Bowden R, Sakaoka H, Donnelly P, Ward R. 2004. High recombination rate in herpes simplex virus type 1 natural populations suggests significant co-infection. Infect Genet Evol 4:115–23.

38. Norberg P, Bergstrom T, Rekabdar E, Lindh M. 2004. Phylogenetic Analysis of Clinical Herpes Simplex Virus Type 1 Isolates Identified Three Genetic Groups and Recombinant Viruses. J Virol 78:10755–10764.

39. Norberg P, Kasubi MJ, Haarr L, Bergstrom T, Liljeqvist J-A. 2007. Divergence and recombination of clinical herpes simplex virus type 2 isolates. J Virol 81:13158–13167.

40. Lonsdale DM, Brown SM, Lang J, Subak-Sharpe JH, Koprowski H, Warren KG. 1980. Variations in herpes simplex virus isolated from human ganglia and a study of clonal variation in HSV-1. Ann N Y Acad Sci 291–308.

41. Liljeqvist JA, Tunback P, Norberg P. 2009. Asymptomatically shed recombinant herpes simplex virus type 1 strains detected in saliva. J Gen Virol 90:559–566.

42. Haarr L, Nilsen A, Knappskog PM, Langeland N. 2014. Stability of glycoprotein gene sequences of herpes simplex virus type 2 from primary to recurrent human infection, and diversity of the sequences among patients attending an STD clinic. BMC Infect Dis 14:63.

43. Kimberlin DW, Lin C-Y, Jacobs RF, Powell DA, Corey L, Gruber WC, Rathore M, Bradley JS, Diaz PS, Kumar M, others. 2001. Safety and efficacy of high-dose intravenous acyclovir in the management of neonatal herpes simplex virus infections. Pediatrics 108:230–238.

44. Kimberlin DW, Whitley RJ, Wan W, Powell DA, Storch G, Ahmed A, Palmer A, Sánchez PJ, Jacobs RF, Bradley JS. 2011. Oral acyclovir suppression and neurodevelopment after neonatal herpes. N Engl J Med 365:1284–1292.

45. Desmyter J, Melnick JL, Rawls WE. 1968. Defectiveness of interferon production and of rubella virus interference in a line of African green monkey kidney cells (Vero). J Virol 2:955–961.

46. Emeny JM, Morgan MJ. 1979. Regulation of the interferon system: evidence that Vero cells have a genetic defect in interferon production. J Gen Virol 43:247–252.

47. Colgrove R, Diaz F, Newman R, Saif S, Shea T, Young S, Henn M, Knipe DM. 2014. Genomic sequences of a low passage herpes simplex virus 2 clinical isolate and its plaque-purified derivative strain. Virology 450–451:140–145.

48. Deschamps T, Kalamvoki M. 2017. Impaired STING Pathway in Human Osteosarcoma U2OS Cells Contributes to the Growth of ICP0-Null Mutant Herpes Simplex Virus. J Virol

91:e00006-17.

49. Yao F, Schaffer PA. 1995. An activity specified by the osteosarcoma line U2OS can substitute functionally for ICP0, a major regulatory protein of herpes simplex virus type 1. J Virol 69:6249–6258.

50. Newman RM, Lamers SL, Weiner B, Ray SC, Colgrove RC, Diaz F, Jing L, Wang K, Saif S, Young S, Henn M, Laeyendecker O, Tobian AAR, Cohen JI, Koelle DM, Quinn TC, Knipe DM. 2015. Genome Sequencing and Analysis of Geographically Diverse Clinical Isolates of Herpes Simplex Virus 2. J Virol JVI.01303-15.

51. Kolb AW, Larsen IV, Cuellar JA, Brandt CR. 2015. Genomic, Phylogenetic, and Recombinational Characterization of Herpes Simplex Virus 2 Strains. J Virol 89:6427–6434.

52. David AT, Saied A, Charles A, Subramanian R, Chouljenko VN, Kousoulas KG. 2012. A Herpes Simplex Virus 1 (McKrae) Mutant Lacking the Glycoprotein K Gene Is Unable To Infect via Neuronal Axons and Egress from Neuronal Cell Bodies. mBio 3.

53. 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.

54. Weber PC, Levine M, Glorioso JC. 1987. Rapid identification of nonessential genes of herpes simplex virus type 1 by Tn5 mutagenesis. Science 236:576–579.

55. Jacobson JG, Martin SL, Coen DM. 1989. A conserved open reading frame that overlaps the herpes simplex virus thymidine kinase gene is important for viral growth in cell culture. J Virol 63:1839–1843.

56. Foster TP, Melancon JM, Baines JD, Kousoulas KG. 2004. The herpes simplex virus type 1 UL20 protein modulates membrane fusion events during cytoplasmic virion morphogenesis and virus-induced cell fusion. J Virol 78:5347–5357.

57. David AT, Baghian A, Foster TP, Chouljenko VN, Kousoulas KG. 2008. The herpes simplex virus type 1 (HSV-1) glycoprotein K(gK) is essential for viral corneal spread and neuroinvasiveness. Curr Eye Res 33:455–467.

58. Dingwell KS, Johnson DC. 1998. The herpes simplex virus gE-gI complex facilitates cell-to-cell spread and binds to components of cell junctions. J Virol 72:8933–8942.

59. Meignier B, Longnecker R, Mavromara-Nazos P, Sears AE, Roizman B. 1988. Virulence of and establishment of latency by genetically engineered deletion mutants of herpes simplex virus 1. Virology 162:251–254.

60. Rochette P-A, Bourget A, Sanabria-Solano C, Lahmidi S, Lavallée GO, Pearson A. 2015. Mutation of UL24 impedes the dissemination of acute herpes simplex virus 1 infection

from the cornea to neurons of trigeminal ganglia. J Gen Virol 96:2794–2805.

61. Jacobson JG, Chen SH, Cook WJ, Kramer MF, Coen DM. 1998. Importance of the herpes simplex virus UL24 gene for productive ganglionic infection in mice. Virology 242:161–169.

62. Leiva-Torres GA, Rochette P-A, Pearson A. 2010. Differential importance of highly conserved residues in UL24 for herpes simplex virus 1 replication in vivo and reactivation. J Gen Virol 91:1109–1116.

63. Liu J, Lewin AS, Tuli SS, Ghivizzani SC, Schultz GS, Bloom DC. 2008. Reduction in severity of a herpes simplex virus type 1 murine infection by treatment with a ribozyme targeting the UL20 gene RNA. J Virol 82:7467–7474.

64. 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 88:1209–27.

65. Pfaff F, Groth M, Sauerbrei A, Zell R. 2016. Genotyping of herpes simplex virus type 1 (HSV-1) by whole genome sequencing. J Gen Virol.

66. Muller W, Jones C, Koelle D. 2010. Immunobiology of Herpes Simplex Virus and Cytomegalovirus Infections of the Fetus and Newborn. Curr Immunol Rev 6:38–55.

67. 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.

68. Snyder A, Polcicova K, Johnson DC. 2008. Herpes simplex virus gE/gI and US9 proteins promote transport of both capsids and virion glycoproteins in neuronal axons. J Virol 82:10613–24.

69. David AT, Baghian A, Foster TP, Chouljenko VN, Kousoulas KG. 2008. The Herpes Simplex Virus Type 1 (HSV-1) Glycoprotein K(gK) is Essential for Viral Corneal Spread and Neuroinvasiveness. Curr Eye Res 33:455–467.

70. Foster TP, Alvarez X, Kousoulas KG. 2003. Plasma membrane topology of syncytial domains of herpes simplex virus type 1 glycoprotein K (gK): the UL20 protein enables cell surface localization of gK but not gK-mediated cell-to-cell fusion. J Virol 77:499–510.

71. 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.

72. Szpara ML, Parsons L, Enquist LW. 2010. Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. J Virol 84:5303–13.

73. Szpara ML, Tafuri YR, Enquist LW. 2011. Preparation of viral DNA from nucleocapsids. J

Vis Exp JoVE 2–7.

74. Szpara M. 2014. Isolation of herpes simplex virus nucleocapsid DNA, p. 31–41. *In* Diefenbach, RJ, Fraefel, C (eds.), Herpes Simplex Virus. Springer New York, New York, NY.

75. 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.

76. Bowen CD, Renner DW, Shreve JT, Tafuri Y, Payne KM, Dix RD, Kinchington PR, Gatherer D, Szpara ML. 2016. Viral forensic genomics reveals the relatedness of classic herpes simplex virus strains KOS, KOS63, and KOS79. Virology 492:179–186.

77. Myers EW, Sutton G, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KHJ, Remington KA, Anson EL, Bolanos RA, Chou H-H, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter CJ. 2000. A Whole-Genome Assembly of Drosophila. Science 287:2196–2204.

78. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol 13:1–9.

79. Warren RL, Sutton GG, Jones SJM, Holt R a. 2007. Assembling millions of short DNA sequences using SSAKE. Bioinformatics 23:500–1.

80. Dolan A, Jamieson FE, Cunningham C, Barnett BC, McGeoch DJ. 1998. The genome sequence of herpes simplex virus type 2. J Virol 72:2010–2021.

81. 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.

82. Strimmer K, Wiuf C, Moulton V. 2001. Recombination analysis using directed graphical models. Mol Biol Evol 18:97–99.

83. Norberg P, Tyler S, Severini A, Whitley R, Liljeqvist J-A, Bergstrom T. 2011. A genome-wide comparative evolutionary analysis of herpes simplex virus type 1 and varicella zoster virus. PloS One 6:1–8.

84. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinforma Appl Note 23:2947–2948.

85. 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.

86. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) 6:80–92.

87. Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X. 2012. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet 3.

88. Brand A, Allen L, Altman M, Hlava M, Scott J. 2015. Beyond authorship: attribution, contribution, collaboration, and credit. Learn Publ 28:151–155.

89. Davison AJ. 2012. Evolution of sexually transmitted and sexually transmissible human herpesviruses. Ann N Y Acad Sci 1230:E37–E49.

90. Lai W, Chen CY, Morse SA, Htun Y, Fehler HG, Liu H, Ballard RC. 2003. Increasing relative prevalence of HSV-2 infection among men with genital ulcers from a mining community in South Africa. Sex Transm Infect 79:202–207.

91. Rawls WE, Laurel D, Melnick JL, Glicksman JM, Kaufman RH. 1968. A search for viruses in smegma, premalignant and early malignant cervical tissues. The isolation of herpesviruses with distinct antigenic properties. Am J Epidemiol 87:647–655.

92. Koelle DM, Norberg P, Fitzgibbon MP, Russell RM, Greninger AL, Huang M-L, Stensland L, Jing L, Magaret AS, Diem K, Selke S, Xie H, Celum C, Lingappa JR, Jerome KR, Wald A, Johnston C. 2017. Worldwide circulation of HSV-2 × HSV-1 recombinant strains. Sci Rep 7:44084.

93. McGeoch DJ, Moss HWM, McNab D, Frame MC. 1987. DNA Sequence and Genetic Content of the HindlIl l Region in the Short Unique Component of the Herpes Simplex Virus Type 2 Genome: Identification of the Gene Encoding Glycoprotein G, and Evolutionary Comparisons. J Gen Virol 68:19–38.

94. Foster TP, Chouljenko VN, Kousoulas KG. 2008. Functional and physical interactions of the herpes simplex virus type 1 UL20 membrane protein with glycoprotein K. J Virol 82:6310–6323.

95. Johnson DC, Frame MC, Ligas MW, Cross AM, Stow ND. 1988. Herpes simplex virus immunoglobulin G Fc receptor activity depends on a complex of two viral glycoproteins, gE and gI. J Virol 62:1347–1354.

96. Lu X, Huang C, Zhang Y, Lin Y, Wang X, Li Q, Liu S, Tang J, Zhou L. 2017. The Us2 Gene Product of Herpes Simplex Virus 2 modulates NF-κB activation by targeting TAK1. Sci Rep 7:8396.

97. Viejo-Borbolla A, Martinez-Martín N, Nel HJ, Rueda P, Martín R, Blanco S, Arenzana-Seisdedos F, Thelen M, Fallon PG, Alcamí A. 2012. Enhancement of chemokine function as an immunomodulatory strategy employed by human herpesviruses. PLoS Pathog

8:e1002497.

98. Cabrera JR, Viejo-Borbolla A, Martinez-Martín N, Blanco S, Wandosell F, Alcamí A. 2015. Secreted herpes simplex virus-2 glycoprotein G modifies NGF-TrkA signaling to attract free nerve endings to the site of infection. PLoS Pathog 11:e1004571.

99. Lymberopoulos MH, Pearson A. 2007. Involvement of UL24 in herpes-simplex-virus-1-induced dispersal of nucleolin. Virology 363:397–409.

100. Foster TP, Melancon JM, Olivier TL, Kousoulas KG. 2004. Herpes simplex virus type 1 glycoprotein K and the UL20 protein are interdependent for intracellular trafficking and trans-Golgi network localization. J Virol 78:13262–13277.