Genetic and phenotypic intrastrain variation in herpes simplex virus type 1 Glasgow strain 17 syn+ derived viruses

Juliet Jones[1], Daniel Pearce Depledge[2#], Judith Breuer[2], Katja Ebert-Keel[1] and Gillian Elliott[1]*

[1] Section of Virology, Department of Microbial Sciences, School of Biosciences and Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK.

[2] Division of Infection & Immunity, University College London, London, UK

[#] Current address: Department of Medicine, New York University School of Medicine, New York, USA.

* Corresponding author. Email: g.elliott@surrey.ac.uk

Keywords: Herpes simplex virus; sequence diversity; cell-to-cell spread; glycoprotein C.

GenBank Accession numbers: s17pp - MN159376; s17pp3 - MN159377; s17pp4 - MN159378; s17pp5 - MN159379; s17pp7a - MN159380; s17pp8 - MN159381; s17pp22a - MN159382; Sc16 - MN159383.

Abstract

The Glasgow s17 syn+ strain of herpes simplex virus 1 (HSV1) is arguably the best characterised strain and has provided the reference sequence for HSV1 genetic studies. Here we show that our original s17 syn+ stock was a mixed population from which we have isolated a minor variant that, unlike other strains in the laboratory, fails to be efficiently released from infected cells and spreads predominantly by direct cell-to-cell transmission. Analysis of other s17-derived viruses that had been isolated elsewhere revealed a number with the same release phenotype. Second generation sequencing of eight plaque-purified s17-derived viruses revealed sequences that vary by 50 SNPs including approximately 10 coding SNPs. This compared to interstrain variations of around 800 SNPs in strain Sc16, of which a quarter were coding changes. Amongst the variations found within s17, we identified thirteen variants of glycoprotein C within the original stock of virus which were predominantly a consequence of altered homopolymeric runs of C residues. Characterisation of seven isolates coding for different forms of gC indicated that all were expressed, despite six of them lacking a transmembrane domain. While the release phenotype did not correlate directly with any of these identified gC variations, further demonstration that nine clinical isolates of HSV1 also fail to spread through extracellular release raises the possibility that propagation in tissue culture had altered the HSV1 s17 transmission phenotype. Hence, this s17 intrastrain variation identified here offers an excellent model for understanding both HSV1 transmission and tissue culture adaptation.

Introduction

Herpes simplex virus 1 (HSV1) infects human mucosal epithelia before establishing lifelong latent infection in sensory neurons [1]. In a proportion of individuals, the virus reactivates periodically to cause recurrent cold sores, genital herpes or other rarer but more serious outcomes including encephalitis and keratitis [2, 3]. As such, the virus has the potential to undergo genetic change within an individual host through sequential episodes of reactivation, or when transmitted from one individual to another.

HSV1 is a large enveloped virus with a 152 kb double-stranded DNA genome coding for at least 75 proteins, which are involved either directly in virus replication, such as for DNA replication or virus assembly, or indirectly in replication through immune evasion [4]. Single or combined genetic changes in these coding regions could dictate variations in the virulence of virus circulating in the human population. The HSV1 genome was first fully sequenced in 1988 when the McGeoch laboratory produced the sequence of the Glasgow strain 17 syn+ virus by extensive DNA fragment cloning and Sanger sequencing [5]. However, it has taken until the last decade and the advent of second generation sequencing for other laboratory strains to be sequenced in full, including KOS [6, 7], McKrae [8] and F [9] allowing sequence comparison between strains that have been propagated in different laboratories. Fewer than 1% single nucleotide polymorphisms (SNPs) have been found between each of these strains and the strain 17 reference sequence, and these are spread across a large number of the 75 virus genes. Roughly one third of SNPs have been seen to be non-synonymous. A smaller number of insertions/deletions (indels) have also been identified in these genomes in comparison to the reference genome, providing scope for a greater degree of divergence. More recently, sequencing of clinical samples of

HSV1 isolated from individuals from a range of global locations, has become a focus of attention for determining the evolution and diversity of HSV1 in the human population [10, 11]. In those studies, specific groups of genes have been determined to be conserved or divergent based on their variability across tens of sequenced genomes, with the highest diversity occuring in a number of envelope glycoproteins [10]. Many viruses were also found to exhibit frameshifts at homopolymeric runs as a result of variation in the length of these sequences [10], extending previous studies that revealed homopolymeric frameshifts in the glycoprotein G (gG) gene of a number of HSV1 and HSV2 clinical isolates [12, 13].

The likelihood of virus adaptation to cell culture may confound the interpretation of sequence differences between clinical and laboratory grown viruses. There is no better example of this than the rapid loss of a large part of the human cytomegalovirus (hCMV) genome over as few as three passages in cultured fibroblasts, resulting in virus that is less cell-associated and reduced in pathogenicity [14]. It is therefore not always clear if specific virus characteristics in culture are relevant to infection in the host. Many of the commonly utilized laboratory strains of HSV1 were isolated decades ago and have been openly distributed to research laboratories. The Glasgow strain 17 syn+ is no exception – originally isolated in 1972 [15], many researchers throughout the world now use it as their prototype virus. By definition however, the strains in use today have an incomplete provenance, as most laboratory strains have been routinely grown in cells of non-human origin - particularly Vero and BHK cells for historical reasons of ease - for an unknown number of passages, and it is possible that genetic adaptation has taken place to such tissue culture conditions. A recent analysis of the KOS strain of HSV1 has looked at the effect of long-term tissue culture passage on that virus and surprisingly found only five coding changes between an

early and recent passage of the virus from the same laboratory [7], while a second study revealed 17 coding changes across three plaque-purified KOS viruses [16]. These studies suggest that HSV1 is relatively stable in culture.

In our laboratory we have routinely used the original Glasgow strain 17 syn+ of HSV1 as both our representative wild-type virus, and the parent virus for production of many recombinant viruses we have generated over the years. Early in our work we took the decision to plaque-purify this virus to take forward as our wild-type into our studies. Here, we show that our plaque purification inadvertently isolated a minor phenotypic variant of Glasgow strain 17 syn+ which has the same plaque morphology as other viruses that were subsequently plaque purified from strain 17 syn+ but has a defective release phenotype compared to the original stock of the virus and laboratory strains Sc16 [17] and KOS. This results in a virus that spreads predominantly cell-to-cell in culture rather than through extracellular release, a phenotype we have now identified in several other strain 17 syn+ derived viruses, most notably the transgenic strain 17 Bac virus. Moreover, sequencing of the viral genome has shown the original Glasgow strain 17 syn+ stock to be a genetically heterogeneous mix of viruses, exemplified by a range of at least 13 glycoprotein C variants including eight frameshift variants. While these gC frameshift variants do not correlate with the identified defective spread phenotype, the further demonstration that nine newly isolated clinical HSV1 viruses were transmitted predominantly by cell-to-cell spread in culture indicates that our laboratory-adapted strains have undergone not only genotypic changes, but phenotypic changes that alter their behaviour in comparison to virus that has been isolated directly from the host. This provides new scope for exploring clinical aspects of HSV1 infection.

Materials and Methods

***Cells and Viruses*.** Vero cells were cultured in DMEM supplemented with 10% newborn calf serum. BSc1, SKOV3 and HCT116 cells were grown in DMEM media supplemented with 10% foetal bovine serum. Wild-type HSV1 Glasgow strain 17 syn+ was kindly provided by Roger Everett, University of Glasgow [15]. Our laboratory strain 17 was subsequently plaque-purified from the original stock of Glasgow strain 17 syn+. Our Δ22 virus [18] is a derivative of the original Glasgow strain 17 syn+ virus, as are the ΔUL13 [19], ΔUL41 [20], and ΔICP0 [21] viruses that were kindly provided by Duncan McGeoch and Roger Everett, University of Glasgow. Transgenic strain 17 (GenBank FJ593289.1) was kindly provided by Andrew Davison, University of Glasgow. Wild-type HSV1 strains Sc16 [17] and KOS [22] were kindly provided by Helena Browne, University of Cambridge and Peter O'Hare, Imperial College London, respectively. All laboratory viruses were routinely propagated and titrated on Vero cells. Clinical variants of HSV1 provided by Mark Atkins, Frimley Park Hospital, were taken as swabs of genital lesions, stored in virus transport medium, and typed on the same day using a real-time PCR commercial kit from Pro Genie Molecular run on a Cepheid Smart Cycler using the manufacturer's cycle conditions. The clinical isolates used in this study were plated out directly using a 1:4 dilution series to determine their plaque morphology without prior propagation in tissue culture.

Virus Spread Assays. Cells grown in six-well plates were infected with approximately 10 to 30 plaque-forming units of virus as determined by titration on Vero cells. Following adsorption for 1 hr, the inoculum was removed and replaced with medium with or without 1% pooled human serum (Seralab). Cells were incubated at 37°C without movement, to ensure only passive transmission of virus through liquid, and at the denoted times removed from the incubator to be fixed and stained with crystal

violet for plaque analysis, or to harvest extracellular virus to determine the amount of virus released.

Illumina Sequencing. Short-read Illumina sequencing was carried out as has been described previously [23]. Sequence datasets were parsed through QUASR [24] and aligned against the s17 reference sequence (GenBank JN555585.1) using BWA [25]. The aligned read data were processed using SAMTools [26] and consensus sequences generated. SNP and indel differences between consensus sequences and the s17 reference sequence were determined using BaseByBase (http://athena.bioc.uvic.ca/) [27].

SDS-PAGE and Western blotting. Protein samples were analysed on 13% polyacrylamide gels and subjected to electrophoresis in Tris-glycine buffer. Gels were transferred to nitrocellulose membrane for Western blot analysis. Western blots were developed using SuperSignal West Pico chemiluminescent substrate (Thermo Fisher Scientific).

Antibodies. Monoclonal anti-VP16 was kindly provided by Colin Crump, University of Cambridge; monoclonal anti-gC and anti-α tubulin were purchased from AbCam and Sigma respectively.

PCR amplification, cloning and sequencing of gC. Virus genomic DNA was extracted from infected cells using a final concentration of 0.5% SDS and 10 mM EDTA, followed by two phenol-chloroform extractions and ethanol precipitation. The gC encoding gene UL44 was amplified by PCR using primers: (5' GGATCCAGGCGTCGGGCATGGCC3') and (5'CTCGAGTTACCGCCGATGACGCTGC 3')

and cloned into the pCR-Blunt-II-TOPO vector (Invitrogen) for downstream sequencing. DNA from plaque purified Glasgow strain 17syn+ viruses was treated in the same way but PCR products were gel purified and sequenced directly. UL25 and UL26 were amplified by PCR using forward and reverse primers (5' GCCGTATGTCCAAGACCTTC 3') and (5' CCCACAGGAATAGGTTGTGG 3'), and (5' CGGGTCTTTCCCAGCATTAC 3') and (5' GTCAACGTCCACGTGTGCT 3') respectively, and sequenced directly. Sequence analysis was carried out using SnapGene software.

Results and Discussion

Laboratory strains of HSV1 exhibit variable ability to spread via an extracellular route. A simple phenotypic assessment of HSV1 fitness is its ability to transmit between cells in culture. The standard method for measuring cell-to-cell spread of many viruses including HSV1 is to conduct a plaque assay whereby released virus is inhibited, either physically by the use of a semi-solid overlay such as agarose or CMC, or immunologically by the inclusion of anti-HSV1 antibodies in the medium overlaying the cell monolayer, thereby measuring direct transfer between cells in contact with each other (cell-to-cell spread in Fig. 1A &1B). In our case, we routinely use 1% pooled human serum to neutralise extracellular virus spread, while allowing cell-to-cell spread, resulting in the production of a plaque that increases in diameter over time. Of note, if the same process is carried out without the inclusion of human serum in the medium, then released virus is free to spread through the extracellular medium, resulting in the formation of characteristic “comet” shaped plaques at early times and total destruction of the monolayer at later times (extracellular spread in Fig. 1A & 1B). Hence the relative size of the comet is an indirect indication of the efficiency of virus release from infected cells, and it is this assay that forms the basis of the study presented here.

While carrying out plaque assays for our lab stock of Glasgow strain 17 syn+ (s17) in parallel to the Sc16 strain, we noted that when human serum was absent from the s17 plaque assays there was little appearance of extracellular spread even at 72h after infection, in contrast to Sc16 where only two plaque-forming units were sufficient to completely destroy the monolayer under the conditions of the experiment here (Fig 1C). Moreover, comparison with the KOS strain indicated that extracellular spread was also a feature of this virus despite it forming smaller plaques overall, suggesting that

the defect in extracellular release of s17 phenotype was unusual (Fig 2A). To quantitate the relative virus release from these three strains, BSC1 cells in six well plates were infected with approximately 30 plaque forming units (PFU) of each strain. At daily intervals, the extracellular media was harvested and the virus in these fractions was titrated. The remaining infected BSC1 cells were fixed and stained to determine the extent of comet tail formation. The stained monolayers confirmed that while both Sc16 and KOS spread efficiently through the extracellular medium causing complete destruction by day 3, s17 formed only localised plaques with most of the monolayer intact at 72 hours (Fig. 2B), with a 1000-fold difference between Sc16 and s17 release from the cell (Fig. 2C).

To ensure this s17 release phenotype was not specific to BSC1 cells, we titrated Sc16 and s17 on to three other cell types – Vero, SKOV3 and HCT116 cells – in both the presence and absence of human serum. In all cases, Sc16 spread extracellularly, but s17 did not (Fig. 3). Moreover, despite the fact that Sc16 formed tiny plaques on HCT116 cells in the presence of human serum, suggesting that the cell-to-cell spread route is inefficient in these cells, the monolayer was completely destroyed in the absence of human serum confirming efficient release and dissemination of Sc16 virions (Fig. 3).

Glasgow Strain17 syn+ contains two release phenotypes. Our s17 virus had been plaque purified from the original strain 17 syn+ virus stock that we had received from Glasgow. This raised the possibility that we had selected a particular subtype of the virus with a different phenotype to the parental virus. Strikingly, comparison of Glasgow strain 17 syn+ (Gla s17) extracellular spread with Sc16, KOS and s17 indicated that while s17 failed to spread further than the cells surrounding the point of infection, the parental Gla s17 virus spread as efficiently as Sc16 and KOS, resulting

230 in complete destruction of the cell monolayer (Fig 4A). These results indicate a genetic
231 variation between our lab s17 and its parent virus that results in a dramatic difference
232 in the ability of the virus to be released from the cell.

233 Due to the fact we had a number of other recombinant viruses in the lab based on
234 Glasgow strain 17 syn+, all of which had by definition been plaque purified during
235 isolation, we next determined the release phenotype of these viruses. This indicated
236 that three of these viruses failed to form comet tails (Fig. 4B, ΔUL13, ΔUL41 and
237 s17Bac virus), while two were released efficiently (Fig. 4B, ΔICP0 and Δ22). Given
238 that gene deletion could potentially confound the interpretation of these results, we
239 next plaque purified a further 18 viruses from the parental Glasgow strain 17 syn+
240 stock and measured their ability to spread by the extracellular route. As shown in Fig
241 5, only one of the 18 viruses (virus 7a) that we selected on that occasion failed to
242 spread extracellularly, confirming that although the phenotype of our original plaque
243 purified s17 had been re-isolated in this experiment, it represented a relatively minor
244 fraction of the parental population. While previous studies have shown that common
245 lab stocks of HSV1 often comprise a minor population of virus with a fusogenic
246 phenotype [16], to our knowledge this is the first time that a virus release phenotype
247 has been identified in this manner. Moreover, given that transgenic s17 also fails to
248 spread extracellularly, it follows that any viruses built in to this backbone are also likely
249 to have this release phenotype.

250 **Clinical isolates of HSV1 spread predominantly by cell-to-cell transmission.**
251 Despite the efficient extracellular release phenotype of the available lab strains, it was
252 not clear if this or the release defective phenotype of s17 would predominate in virus
253 causing clinical infection. Hence, to assess the spread phenotypes of several
254 representative clinical HSV1 isolates, we carried out the same extracellular and cell-

255 to-cell spread assays on nine newly obtained clinical samples of HSV1. In contrast to
256 the efficiently released Sc16, there was little evidence of spread by virus release in
257 any of these clinically isolated viruses (Fig 6). This therefore raises the question of
258 whether active release of virions from the cell surface is a natural transmission route
259 for HSV1 in the infected host [28]. Other alphaherpesviruses including varicella zoster
260 virus and Marek's disease virus are known to be highly cell-associated in culture [29,
261 30]. Moreover, clinical isolates of hCMV are initially cell-associated but have been
262 shown to mutate rapidly in cell culture changing to a virus that is efficiently released
263 [14]. While some viruses such as influenza virus use neuraminidase in the envelope to
264 actively promote virion release into the extracellular environment by cleavage of the
265 glycoprotein attachment site followed by transmission to distal locations [31], other
266 viruses have evolved to spread mainly by direct cell-to-cell transfer, whereby the virus
267 particle remains tethered to the cell surface and can only spread into a cell directly in
268 contact with the infected cell, protecting it from recognition by the immune system. The
269 classical example of such a virus is HIV, which transmits between T cells via the
270 virological synapse [32]. Although the molecular detail of HSV1 transmission in the
271 host is poorly characterized, given the phenotypes of other herpesviruses and our
272 data presented here, it is possible that commonly used lab strains of HSV1 have
273 undergone genetic changes that alter a predominantly cell-associated phenotype to
274 one that, similar to hCMV, becomes more actively released from the cell. However, it
275 should be noted that recent studies on ten random clinical HSV1 strains isolated in
276 Finland indicated that while these viruses replicated with lower efficiencies than strain
277 17 syn+ in all cell types tested, and therefore also released fewer virions, the overall
278 proportion of released virus was similar and in some cases higher than strain 17 syn+
279 [33]. While the nature of the assays carried out in those studies was different to our

280 own (high multiplicity versus spread at low multiplicity respectively), further studies will
281 be required to determine the relationship between efficient virus release and
282 propagation in cell culture. Nonetheless, our results suggest that it would be prudent
283 to combine studies of more clinically relevant HSV1 strains with those of common lab
284 strains when investigating HSV1 transmission mechanisms. Moreover, given that
285 HSV1 infects stratified differentiated epithelia in the host, rather than monolayers as
286 used here, it will be important to determine how the release phenotype affects virus
287 spread in systems such as organotypic raft culture or human skin explant.

288 **Genome sequencing of strain 17 plaque purified viruses.** The above results
289 indicated that we had isolated two phenotypes from the starting Glasgow strain 17
290 syn+ virus stock. To determine if we could identify common genetic difference(s) in the
291 three s17 viruses that fail to release virus efficiently, we carried out second generation
292 sequencing of seven plaque-purified s17 genomes together with our own lab stock of
293 the transgenic s17 (s17Bac), totalling four that release virus and three that are
294 defective. In addition, we sequenced our lab stock of Sc16 genome to compare
295 relative intrastrain variation to interstrain variation of HSV1 grown in our laboratory.
296 While sequencing of the Sc16 genome revealed around 800 SNPs of which ~190
297 were non-synonymous to the strain 17 reference sequence (JN555585.1), in broad
298 agreement with the published sequence for Sc16 (GenBank KX946970.1) [34], only
299 around 50 SNPs were found for each of the eight representative strain 17 genomes
300 that we sequenced, including 6 to 15 non-synonymous changes in each virus (Table 1
301 & Fig. 7). These statistics are in line with previously published intrastrain sequencing
302 variations from single virus stocks for three isolates of strain KOS and two isolates of
303 strain F [16]. Apart from inaccuracies in the repeat regions of the genome, our
304 transgenic s17 sequence matched that for transgenic strain 17 in Genbank

(FJ593289.1), confirming that our sequencing pipeline was accurate. Moreover, our Sc16 sequence revealed ony 4 non-synonymous SNPs compared to the published sequence [34], despite each sequenced stock being propagated in different laboratories. Interestingly, of the identified coding changes in the sequenced s17 genomes, only three were present in the three non-spreading viruses – in genes UL25, UL26 and UL55 – but not in the other five viruses (in bold, Table 1; Fig. 7). Sanger sequencing of the UL25 and UL26 genes confirmed that those two variations also segregated appropriately in the Δ41 and ΔUL13 viruses which had been isolated several decades ago in a different laboratory [19, 20], and which as shown above also failed to spread extracellularly (Fig 4B). Such analyses do not prove that either or both of these variations are responsible for the defective release phenotype of these viruses, but it is interesting to note that UL26 encodes a protease which although characterized to be involved in capsid assembly, may nonetheless has the potential to have other activities during infection such as facilitating virus release from the cell surface. Although it has not yet been possible to directly correlate intrastrain genetic changes with phenotypic changes in these viruses, it should be noted that a recent study of clinical isolates also failed to identify sequences that would be predictive for phenotypic differences of virus replication and release [33], suggesting that these phenotypes are complex and multifactorial.

Hypervariable gC sequences in Glasgow strain 17 syn+. In the course of this genomic sequencing we identified another frequent S57N variation in the gC open reading frame, which despite its possible relevance to a virus release phenotype, was not seen to segregate with either the positive or negative release phenotypes (Table 1). Nonetheless, we also identified a potential single nucleotide deletion at a homopolymeric run at proline residue 196 in the gC gene of pp3 and pp7a viruses,

which if correct would result in a frameshift and premature termination at residue 198 of the open reading frame. This was similar to a previously identified variant of gC in the genome of the Glasgow strain 17 syn+ - derived ICP0 deletion mutant dl1043 [35], where a similar single nucleotide deletion at a homopolymeric run at proline 62 followed by termination at 175 was identified [36], suggesting that the gC gene is susceptible to such variations. We carried out Sanger sequencing of the gC gene from the 18 plaque-purified viruses denoted in Fig 5. Within these viruses we found the reference sequence, and another seven variants of gC, including the S57N variation, five which contained different frameshifts and premature stop codons, and one which contained a premature stop (Fig 8A). To confirm expression of each of these gC proteins, we analysed cell lysates from cells infected with these viruses by SDS-PAGE and Western blotting for gC (Fig 8B). This indicated that all viruses expressed their gC proteins, and those viruses containing premature stop codons produced significantly smaller gC peptides. Furthermore, the fact that this monoclonal antibody recognised every gC variant placed its epitope at the N-terminus of the protein. With the exception of the shortest molecule, which migrated at only a slightly higher molecular weight than predicted, the sizes of the other gC molecules were larger than their primary sequences predicted (Fig 8B), indicating that these gC variants were glycosylated in a manner supported by previously mapped glycosylation sites (Fig 8A). Although gC in pp22a was full-length, it was consistently expressed at a very low level, suggesting that another indirect mechanism such as mutation of the gene promoter may be responsible. Further PCR, cloning and sequencing of thirty gC clones from the original strain 17 virus stock revealed yet another four variants of gC (Table 2).

In total, we have identified 13 variants of gC in the original stock of Glasgow strain 17 syn+, extending a previous study that had already identified one of these [36]. While it

355 remains to be determined why our original strain 17 stock contained quite so many gC
356 variants, we suggest that these variants could be present in many strain 17 derived
357 viruses dispersed throughout the world. Eight of these variants involve a one
358 nucleotide frameshift at homopolymeric G/C tracts, resulting in a range of prematurely
359 terminated proteins that all lack the gC transmembrane domain, producing proteins
360 that as we have shown above are expressed and in theory would be secreted rather
361 than incorporated into the virus envelope. A similar frameshift has been reported
362 before in the macroplaque strain of HSV1 which also fails to express full-length gC
363 [37]. Splicing of the gC mRNA has also been shown to result in a secreted form of gC
364 in infected cells due to the lack of a transmembrane domain [38], with the authors of
365 that study suggesting that this form of secreted gC could act as a virulence factor. Of
366 note, the varicella zoster virus glycoprotein C has also been shown to be
367 hypervariable due to strain variation in copy numbers of a repeat sequence within the
368 gene [39], suggesting that gC variation may be a common alphaherpesvirus trait.
369 Another example of an HSV1 gene that develops variation due to insertions or
370 deletions at homopolymeric runs is the thymidine kinase gene, particularly in patients
371 in response to antiviral treatment with acyclovir [40]. Glycoprotein G has also been
372 shown to be highly variable, with around 40% of isolates exhibiting genetic variation in
373 one study [12]. Amongst these variations within gG, homopolymer frameshifts have
374 been identified in a small number of natural clinical isolates of HSV1 and HSV2 [12,
375 13]. What is striking about our strain 17 syn+ gC result presented here is that all 13 gC
376 variations were found in a single aliquot of the original virus stock received from
377 Glasgow, indicating that it already contained this highly mixed genomic population.
378 The mechanism of insertion or deletion at such homopolymeric runs is likely to be due
379 to the HSV1 polymerase slipping or stuttering in a similar manner to that found to

occur in the mRNA editing mechanism of paramyxovirus RNA polymerases, which insert a variable number of G residues at a specific homopolymeric run within their P genes [41]. Given the extent of homopolymeric runs distributed throughout the HSV1 genome [10], it is highly likely that further diversity in these genes will emerge as more sequences are obtained from isolates obtained from across the globe. In summary, our results presented here add to the growing understanding of HSV1 genetic diversity and evolution in both the clinic and the laboratory, and have potential implications for studies involving strain 17-derived viruses.

Author Contributions

JJ: investigation; validation

DPD: formal analysis; data curation

JB: resources

KEK: investigation

GE: conceptualisation; investigation; writing – original draft preparation; writing – review and editing; visualisation; supervision; funding.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest

Funding Information

This work was funded by Worldwide Cancer Research grant 15-0341 to GE.

Acknowledgements

We thank Mark Atkins (Frimley Park Hospital) for the clinical HSV1 isolates used in this study. We also thank Roger Everett, Andrew Davison, Helena Browne and Peter O'Hare for the kind donation of other viruses.

References

1. **Wilson AC, Mohr I**. A cultured affair: HSV latency and reactivation in neurons. *Trends Microbiol* 2012;20(12):604-611.

2. **Whitley RJ**. Herpes simplex encephalitis: adolescents and adults. *Antiviral Res* 2006;71(2-3):141-148.

3. **Farooq AV, Shukla D**. Herpes simplex epithelial and stromal keratitis: an epidemiologic update. *Survey of ophthalmology* 2012;57(5):448-462.

4. **Tognarelli EI, Palomino TF, Corrales N, Bueno SM, Kalergis AM et al.** Herpes simplex virus evasion of early host antiviral responses. *Front Cell Infect Microbiol* 2019;9:127.

5. **McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC et al.** The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J Gen Virol* 1988;69(1531):1531-1574.

6. **Macdonald SJ, Mostafa HH, Morrison LA, Davido DJ**. Genome sequence of herpes simplex virus 1 strain KOS. *J Virol* 2012;86(11):6371-6372.

7. **Colgrove RC, Liu X, Griffiths A, Raja P, Deluca NA et al.** History and genomic sequence analysis of the herpes simplex virus 1 KOS and KOS1.1 sub-strains. *Virology* 2016;487:215-221.

8. **Macdonald SJ, Mostafa HH, Morrison LA, Davido DJ**. Genome sequence of herpes simplex virus 1 strain McKrae. *J Virol* 2012;86(17):9540-9541.

9. **Szpara ML, Parsons L, Enquist LW**. Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. *J Virol* 2010;84(10):5303-5313.

10. **Szpara ML, Gatherer D, Ochoa A, Greenbaum B, Dolan A et al.** Evolution and diversity in human herpes simplex virus genomes. *J Virol* 2014;88(2):1209-1227.

11. **Pfaff F, Groth M, Sauerbrei A, Zell R**. Genotyping of herpes simplex virus type 1 by whole-genome sequencing. *J Gen Virol* 2016;97(10):2732-2741.

12. **Rekabdar E, Tunback P, Liljeqvist JA, Lindh M, Bergstrom T**. Dichotomy of glycoprotein g gene in herpes simplex virus type 1 isolates. *J Clin Microbiol* 2002;40(9):3245-3251.

13. **Liljeqvist JA, Svennerholm B, Bergstrom T**. Herpes simplex virus type 2 glycoprotein G-negative clinical isolates are generated by single frameshift mutations. *J Virol* 1999;73(12):9796-9802.

14. **Wilkinson GW, Davison AJ, Tomasec P, Fielding CA, Aicheler R et al.** Human cytomegalovirus: taking the strain. *Medical microbiology and immunology* 2015;204(3):273-284.

15. **Brown SM, Ritchie DA, Subak-Sharpe JH**. Genetic studies with herpes simplex virus type 1. The isolation of temperature-sensitive mutants, their arrangement into complementation groups and recombination analysis leading to a linkage map. *J Gen Virol* 1973;18(3):329-346.

16. **Parsons LR, Tafuri YR, Shreve JT, Bowen CD, Shipley MM et al.** Rapid genome assembly and comparison decode intrastrain variation in human alphaherpesviruses. *MBio* 2015;6(2).

17. **Hill TJ, Field HJ, Blyth WA**. Acute and recurrent infection with herpes simplex virus in the mouse: a model for studying latency and recurrent disease. *J Gen Virol* 1975;28(3):341-353.

18. **Elliott G, Hafezi W, Whiteley A, Bernard E**. Deletion of the herpes simplex virus VP22-encoding gene (UL49) alters the expression, localization, and virion incorporation of ICP0. *J Virol* 2005;79(15):9735-9745.

19. **Coulter LJ, Moss HW, Lang J, McGeoch DJ**. A mutant of herpes simplex virus type 1 in which the UL13 protein kinase gene is disrupted. *J Gen Virol* 1993;74 (387-395.

20. **Fenwick ML, Everett RD**. Inactivation of the shutoff gene (UL41) of herpes simplex virus types 1 and 2. *J Gen Virol* 1990;71 2961-2967.

21. **Everett RD, Boutell C, Orr A**. Phenotype of a herpes simplex virus type 1 mutant that fails to express immediate-early regulatory protein ICP0. *J Virol* 2004;78(4):1763-1774.

22. **Smith KO**. Relationship between the Envelope and the Infectivity of Herpes Simplex Virus. *Proc Soc Exp Biol Med* 1964;115:814-816.

23. **Milne I, Bayer M, Cardle L, Shaw P, Stephen G et al.** Tablet--next generation sequence assembly visualization. *Bioinformatics* 2010;26(3):401-402.

24. **Watson SJ, Welkers MR, Depledge DP, Coulter E, Breuer JM et al.** Viral population analysis and minority-variant detection using short read next-generation sequencing. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2013;368(1614):20120205.

25. **Li H, Durbin R**. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25(14):1754-1760.

26. **Bellastella A, Amato G, Bizzaro A, Carella C, Criscuolo T et al.** Light, blindness and endocrine secretions. *Journal of endocrinological investigation*, Review 1999;22(11):874-885.

27. **Hillary W, Lin SH, Upton C**. Base-By-Base version 2: single nucleotide-level analysis of whole viral genome alignments. *Microb Inform Exp* 2011;1(1):2.

28. **Sattentau Q**. Avoiding the void: cell-to-cell spread of human viruses. *Nat Rev Microbiol* 2008;6(11):815-826.

29. **Weller TH, Witton HM**. The etiologic agents of varicella and herpes zoster; serologic studies with the viruses as propagated in vitro. *The Journal of experimental medicine* 1958;108(6):869-890.

30. **Biggs PM**. The history and biology of Marek's disease virus. *Curr Top Microbiol Immunol* 2001;255:1-24.

31. **Dou D, Revol R, Ostbye H, Wang H, Daniels R**. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. *Front Immunol* 2018;9:1581.

32. **Jolly C, Kashefi K, Hollinshead M, Sattentau QJ**. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. *The Journal of experimental medicine* 2004;199(2):283-293.

33. **Bowen CD, Paavilainen H, Renner DW, Palomaki J, Lehtinen J et al.** Comparison of herpes simplex virus 1 strains circulating in Finland demonstrates the uncoupling of whole-genome relatedness and phenotypic outcomes of viral infection. *J Virol* 2019;93(8).

34. **Rastrojo A, Lopez-Munoz AD, Alcami A**. Genome Sequence of Herpes Simplex Virus 1 Strain SC16. *Genome Announc* 2017;5(4).

35. **Stow ND, Stow EC**. Isolation and characterization of a herpes simplex virus type 1 mutant containing a deletion within the gene encoding the immediate early polypeptide Vmw110. *J Gen Virol* 1986;67:2571-2585.

36. **Cunha CW, Taylor KE, Pritchard SM, Delboy MG, Komala Sari T et al.** Widely used herpes simplex virus 1 ICP0 deletion mutant strain dl1403 and Its derivative viruses do not express glycoprotein C due to a secondary mutation in the gC gene. *PloS one* 2015;10(7):e0131129.

37. **Draper KG, Costa RH, Lee GT, Spear PG, Wagner EK**. Molecular basis of the glycoprotein-C-negative phenotype of herpes simplex virus type 1 macroplaque strain. *J Virol* 1984;51(3):578-585.

38. **Sedlackova L, Perkins KD, Lengyel J, Strain AK, van Santen VL et al.** Herpes simplex virus type 1 ICP27 regulates expression of a variant, secreted form of glycoprotein C by an intron retention mechanism. *J Virol* 2008;82(15):7443-7455.

39. **Kinchington PR, Remenick J, Ostrove JM, Straus SE, Ruyechan WT et al.** Putative glycoprotein gene of varicella-zoster virus with variable copy numbers of a 42-base-pair repeat sequence has homology to herpes simplex virus glycoprotein C. *J Virol* 1986;59(3):660-668.

40. **Gaudreau A, Hill E, Balfour HHJ, Erice A, Boivin G**. Phenotypic and genotypic characterization of acyclovir-resistant herpes simplex viruses from immunocompromised patients. *J Inf Dis* 1998;178(2).

41. **Kolakofsky D**. Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: A review. *Virology* 2016;498:94-98.

s17 Ref	Gene	*s17Bac*	*s17*	pp3	pp4	pp5	pp8	*pp7a*	pp22a
3475	RL2					R150H			
3497	RL2				N157K		N157K	N157K	
4370	RL2				A403P				
4372	RL2								A404P
4374	RL2				V405A	V405A			
4383	RL2					C408S			
5113	RL2						V651M		
9978	UL1						S214N		
9978	UL2								
10720	UL2								G279D
11495	UL3						R169*		
12525	UL5						L870I		
14214	UL5			L307M					
20808	UL9	S818N							
24039	UL10	V279M						V279M	
	UL11	E39K			E39K		E39K	E39K	
26671	UL12		R73H				R73H	R73H	
26865	UL12		C9Y				C9Y	C9Y	
26866	UL12	A8V							
27473	UL13	T344I							
36965	UL19	A1189T							
42783	UL21				A237T				
45357	UL22					G343C			
45656	UL22				T243M				T243M
46927	UL23	R293W						R293W	
47909	UL24					V58I			
48533	UL24			A266T					
49496	**UL25**	**H228R**	**H228R**					**H228R**	
51314	UL26								D169N
52476	**UL26**	**T556I**	**T556I**					**T556I**	
52476	**UL26.5**	**T250I**	**T250I**					**T250I**	
60496	UL29					A520V			
60874	UL29							A394V	
65210	UL30			L802F					
66850	UL31								A177V
69041	UL32								A41V
69646	UL34						G5D		
70353	UL34					R241W			
71581	UL36				E2963A				
72307	UL36							P2721H	
76283	UL36			L1396F		L1396F			
78550	UL36				A640V				
92332	UL41								T102M
94405	UL42								D432N
96481	UL44	S57N	S57N			S57N	S57N		S57N

99225	UL46								A577T
100833	UL46				L41F				
100949	UL46	R3H	R3H			R3H			
104378	UL48	T235A							
104966	UL48					A39T			
105735	UL49	L220I							
106334	UL49	E20G							
106750	UL49A								C82Y
107119	UL50							Q37*	
107228	UL50						A73V		
114857	UL54		R374C						
115516	UL55							T7I	
115918	UL55								C141Y
115922	**UL55**	**H142Q**	**H142Q**					**H142Q**	
115952	UL55			W152*	W152*		W152*		
128783	RS1					P783R			
128786	RS1					Q782G			
128790	RS1					A781I			
128792	RS1					P780Q			
133070	US1						V142F		
133713	US1	DS356G							
134133	US2					E267K			
139897	US7					D37Y			
139903	US7			G39R	G39R		G39R		
140564	US7								T259M
144820	US10						S94N	S94N	
144873	US10		N76K						
144873	US11		P127T						
144894	US11								D120N
Coding changes		17	11	6	11	18	17	16	14
Total SNPs		27	22	47	56	61	39	50	55

Table 1: Non-synonymous changes in viruses derived from Glasgow strain 17 syn+. Genomic sequences of seven plaque-purified viruses were aligned with the strain 17 reference sequence (JN555585.1). BAC refers to transgenic strain 17 sequenced from the transgenic strain 17 virus that had been grown in our laboratory. SNPs shared between the three viruses defective for extracellular release (s17Bac, s17 and pp7a) are in bold.

Coding Change	Frameshift	Length (aa)	Plaques (18)	Clones (26)
		511	1	5
S57N		511	3	3
S57N/R155L		511	0	1
	P371	409	0	1
Q307*		307	1	0
S57N	H202	229	1	0
	P196	198	4	6
	G188	198	1	0
	P122	175	0	1
S57N	P122	175	0	1
	P88	175	0	1
S57N	P88	175	1	1
S57N	P62	175	6	6

Table 2: Glycoprotein C variants found in Glasgow strain 17 syn+. The gC gene was sequenced by PCR amplification of 18 plaque purified viruses and 26 clones of PCR amplified gC DNA from the original strain 17 syn+ virus stock.

Figure Legends

Figure 1: Differential release phenotypes of two laboratory strains of HSV1. (A) HSV1 can spread by direct cell-to-cell spread or extracellular release followed by spread to distal cells. Direct spread is assessed by performing plaque assays in the presence of neutralizing human serum; extracellular spread is assessed by performing plaque assays in the absence of human serum. (B) Approximately 10 pfu of HSV1 (strain Sc16) was plated on to BSc1 cells and incubated in the presence or absence of 1% human serum to inhibit extracellular virus. (C) Strains Sc16 and s17 were titrated onto BSc1 cells in the presence (cell-to-cell) or absence (extracellular) of human serum. Plates were fixed and stained with crystal violet 36 and 72 hours later.

Figure 2: Strain 17 is defective for extracellular release. (A) Strain 17, Sc16 and KOS strains of HSV1 were titrated on to BSc1 cells in the presence or absence of human serum and fixed and stained with crystal violet 48 hours later. (B) & (C) Approximately 30 plaque forming units of Sc16, KOS or strain 17 virus was used to infect BSc1 cells grown in a six well plate. At daily intervals the supernatant was harvested and titrated on to Vero cells (C) and the remaining cells fixed and stained with crystal violet.

Figure 3: Strain 17 is defective for extracellular release in a range of cell types. Strain 17 and Sc16 strains of HSV1 were titrated on to Vero, SKOV3 and HCT116 cells in the presence or absence of human serum and fixed and stained with crystal violet 72 hours later.

Figure 4: The original Glasgow strain 17 syn+ stock contained two extracellular release phenotypes. (A) BSc1 cells grown in 6-well plates were infected with approximately 30 plaque forming units of Sc16, KOS, Glasgow strain 17 and plaque purified s17 in the presence or absence of human serum. Cells were fixed at 72 hours

564 and stained with crystal violet. (B) As in A, but cells were infected with strain 17-
565 derived viruses as denoted.

566 **Figure 5: Extracellular release phenotypes of multiple plaque purified Glasgow**
567 **strain 17 isolates.** Eighteen plaques were purified from Glasgow strain 17 syn+ virus,
568 and approximately 10 plaque forming units of each was plated onto BSc1 cells grown
569 in 6-well plates in the absence of human serum. The cells were fixed after 72 hours
570 and stained with crystal violet.

571 **Figure 6: Clinical isolates of HSV1 exhibit defective extracellular release.** Nine
572 clinical isolates were plated onto BSc1 cells grown in 6-well plates in the absence or
573 presence of human serum, alongside Sc16. The cells were fixed after 72 hours and
574 stained with crystal violet.

575 **Figure 7: Single nucleotide polymorphisms (SNPs) in seven strain 17 syn+**
576 **derived isolates.** The HSV-1 genome is represented in yellow with ORFs arrayed
577 above and below (dark grey) according to strand. Each light grey band indicates a
578 different sequenced genome, track order (out to in) is
579 s17>s17Bac>pp7a>pp3>pp4>pp5>pp8>pp22a. Red indicates non-synonymous
580 SNPs and black indicates non-coding or synonymous SNPs. Coding SNPs common to
581 the three release defective viruses (s17, s17Bac and pp7a) are shown in blue, as are
582 the ORFs in which they are located (UL25, UL26 and UL55). The gC ORF (UL44) has
583 also been highlighted in blue (also see Fig 8).

584 **Figure 8: Glasgow strain 17 syn+ contains multiple variants of glycoprotein C.**
585 (A) Eight variants of gC were identified in the strain 17 plaques shown in Fig 5,
586 including the original reference sequence. Six of these contained frameshifts at
587 homopolymeric runs resulting in proteins truncated at premature stop codons.

588 Diamonds denote known glycosylation sites. The predicted and actual molecular
589 weight of these molecules (as identified in B) are shown on the right-hand side. (B)
590 BSc1 cells infected with the denoted viruses were harvested at 18 hours and analysed
591 by SDS-PAGE on a 13% polyacrylamide gel followed by Western blotting for gC,
592 VP16 and α-tubulin. Molecular weight markers are shown on right-hand side in kDa.
593
594

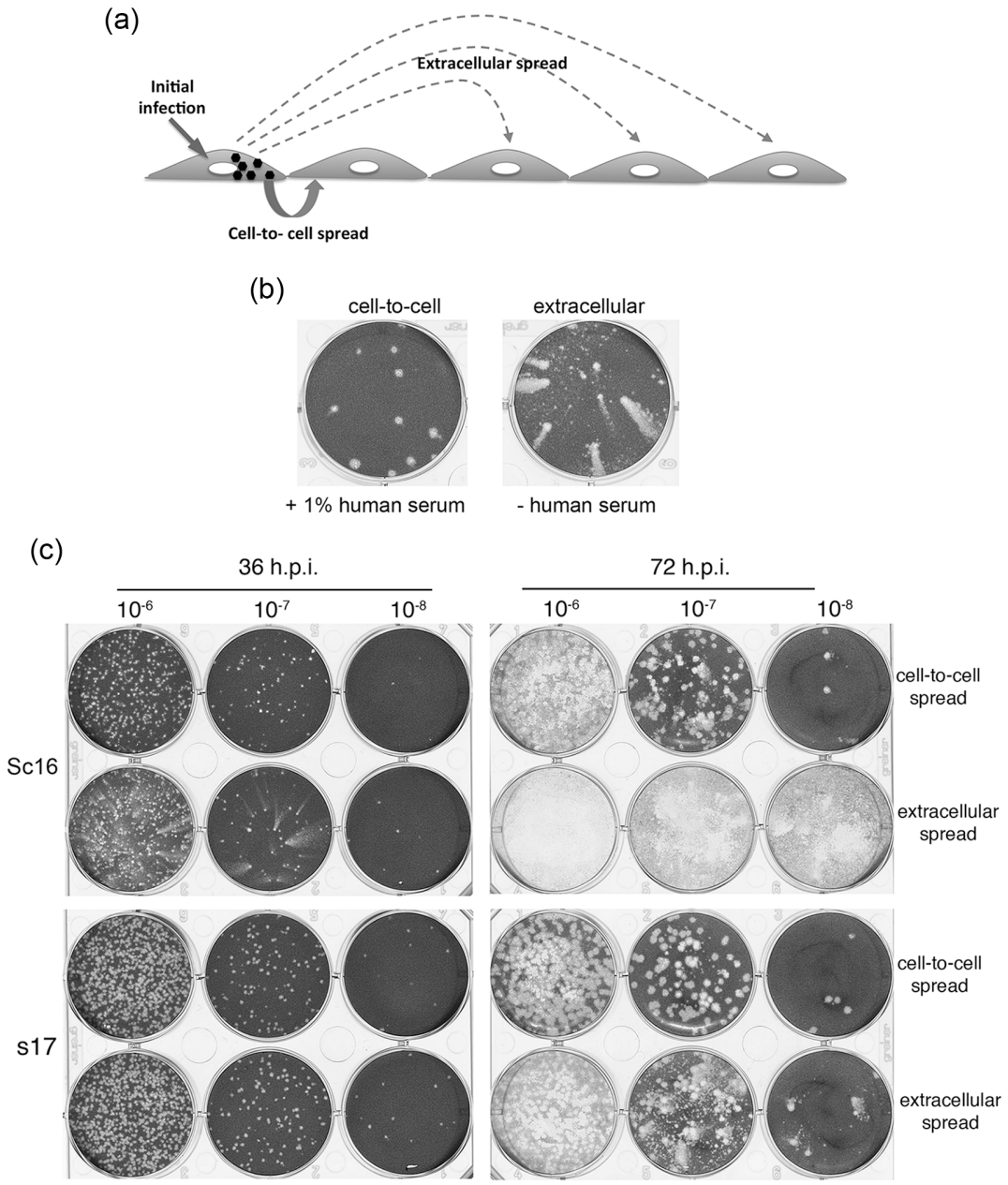

(a)
Initial infection
Extracellular spread
Cell-to- cell spread
(b)
cell-to-cell
extracellular
+ 1% human serum
- human serum
(c)
36 h.p.i.
72 h.p.i.
10-6
10-7
10-8
10-6
10-7
10-8
Sc16
cell-to-cell spread
extracellular spread
s17
cell-to-cell spread
extracellular spread

Figure 1

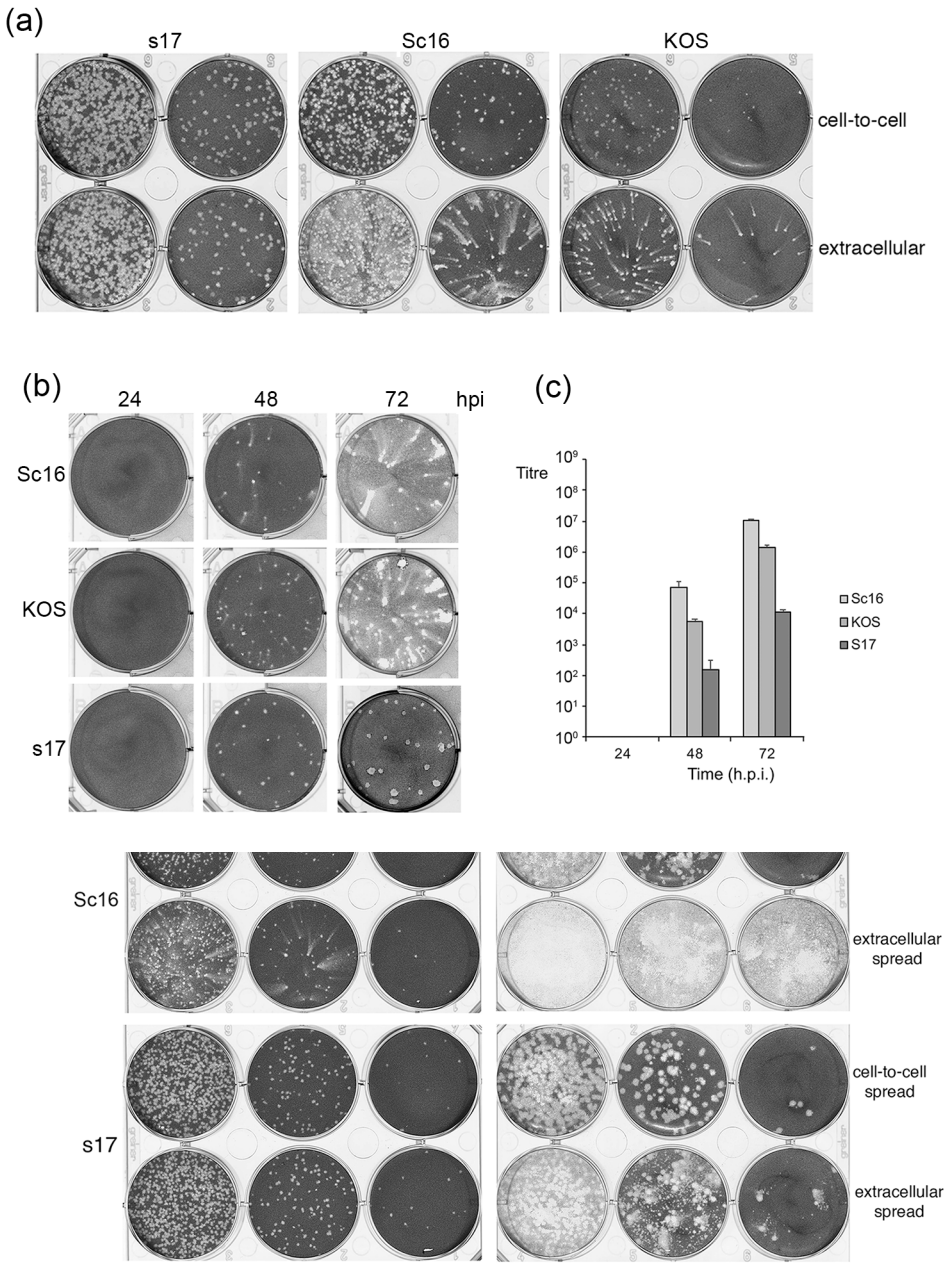

(a)
s17
Sc16
KOS
cell-to-cell
extracellular
(b)
24
48
72
hpi
Sc16
KOS
s17
(c)
Titre
Sc16
KOS
S17
Time (h.p.i.)
Sc16
extracellular spread
cell-to-cell spread
s17
extracellular spread

Figure 2

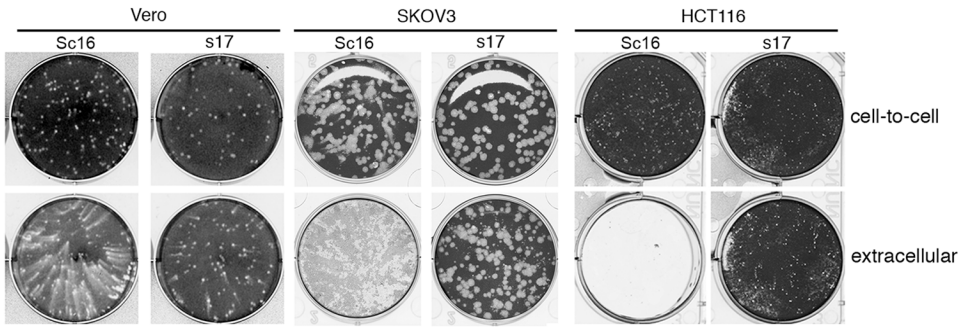
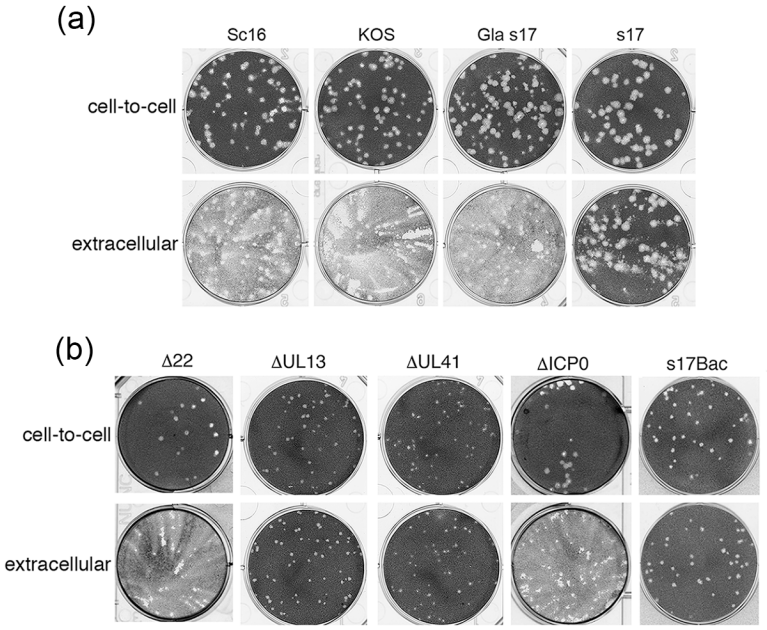

Vero
SKOV3
HCT116
Sc16
s17
Sc16
s17
Sc16
s17
cell-to-cell
extracellular

Figure 3

610



611

612 **Figure 4**

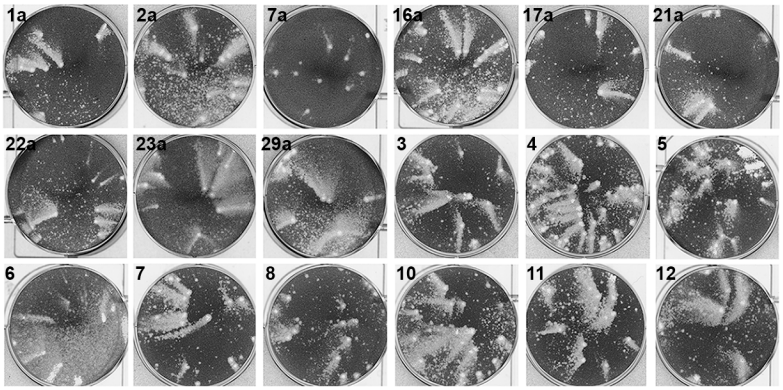
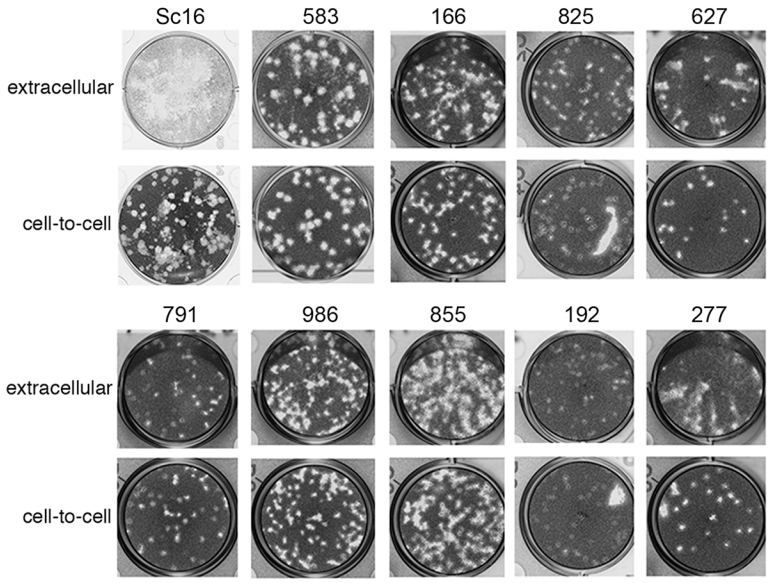


Figure 5

620

621



622

623 **Figure 6**

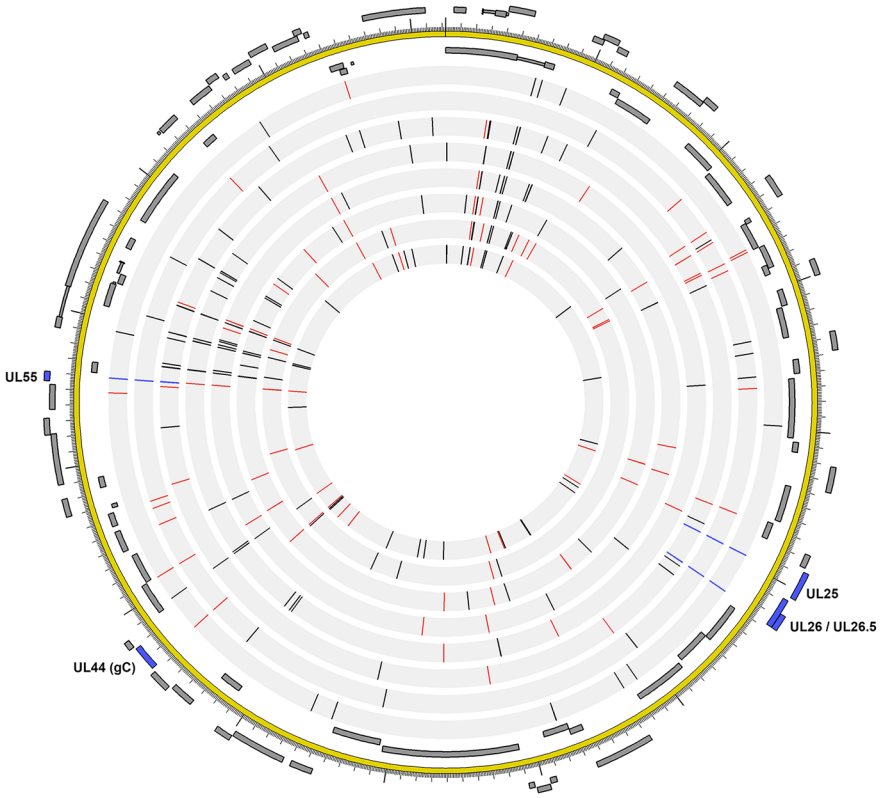


Figure 7

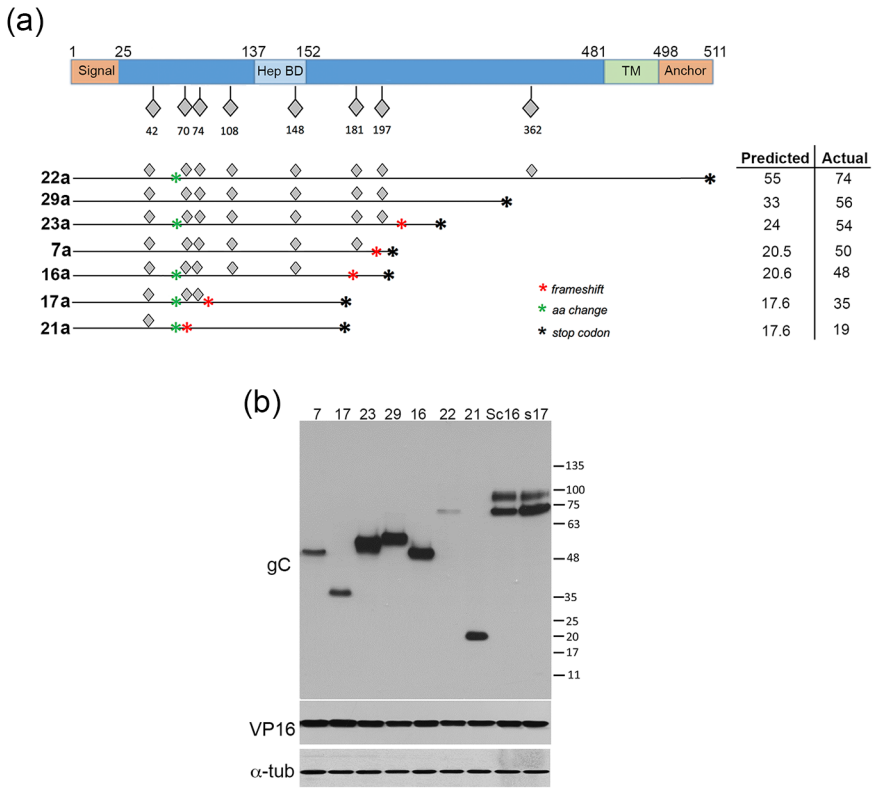

(a)
1
25
137
152
481
498
511
Signal
Hep BD
TM
Anchor
42
70 74
108
148
181
197
362
22a
29a
23a
7a
16a
17a
21a
* frameshift
* aa change
* stop codon
Predicted
Actual
55
74
33
56
24
54
20.5
50
20.6
48
17.6
35
17.6
19
(b)
7
17
23
29
16
22
21
Sc16
s17
135
100
75
63
48
35
25
20
17
11
gC
VP16
α-tub

Figure 8