Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication


1Section of Medical Microbiology, Tenovus Building, School of Medicine, Cardiff University, Cardiff, United Kingdom. 2MRC Virology Unit, Institute of Virology, University of Glasgow, Church Street, Glasgow, United Kingdom.

Human cytomegalovirus (HCMV) in clinical material cannot replicate efficiently in vitro until it has adapted by mutation. Consequently, wild-type HCMV differ fundamentally from the passaged strains used for research. To generate a genetically intact source of HCMV, we cloned strain Merlin into a self-excising BAC. The Merlin BAC clone had mutations in the RL13 gene and UL128 locus that were acquired during limited replication in vitro prior to cloning. The complete wild-type HCMV gene complement was reconstructed by reference to the original clinical sample. Characterization of viruses generated from repaired BACs revealed that RL13 efficiently repressed HCMV replication in multiple cell types; moreover, RL13 mutants rapidly and reproducibly emerged in transfectants. Virus also acquired mutations in genes UL128, UL130, or UL131A, which inhibited virus growth specifically in fibroblast cells in wild-type form. We further report that RL13 encodes a highly glycosylated virion envelope protein and thus has the potential to modulate tropism. To overcome rapid emergence of mutations in genetically intact HCMV, we developed a system in which RL13 and UL131A were conditionally repressed during virus propagation. This technological advance now permits studies to be undertaken with a clonal, characterized HCMV strain containing the complete wild-type gene complement and promises to enhance the clinical relevance of fundamental research on HCMV.

Introduction

Human cytomegalovirus (HCMV) is a clinically important herpesvirus that is ubiquitous in human populations worldwide (1). Primary infection is followed by lifelong persistence, during which virus reactivation must be constrained continuously by host immune surveillance. Myeloid cell progenitors are a recognized site of latency, with infectious virus being produced following differentiation into macrophages or dendritic cells. Severe disease is most commonly observed when immunity is compromised by infection (e.g., HIV/AIDS) or immunosuppressive therapy (e.g., in transplant recipients). HCMV is also the leading viral cause of congenital disability and malformation, which was the primary basis for it being designated a highest priority level vaccine target by the US Institute of Medicine (2). HCMV disease can manifest as a wide range of clinical conditions (e.g., pneumonia, colitis, retinitis, hepatitis, arteriosclerosis, or systemic infection), reflecting the capacity of the virus to infect a wide range of cell types in vivo (3–7). Despite this wide tropism in vivo, only fibroblasts support the efficient growth of cultured strains in vitro.

Appreciation of HCMV pathogenesis is heavily dependent on research performed using the high-passage strains AD169 and Towne. However, these strains are known to have both lost virulence and experienced substantial alterations in their genomes during passage in cell culture (8–10). The HCMV genome may be represented as ab-U1-b′a′c′-U2-ca, where U1 and U2 denote long and short unique regions and ba/b′a′ and ca/c′a′ indicate inverted repeats flanking the unique regions. AD169 and Towne have acquired large deletions (15 kb and 13 kb, respectively) in the U1′/b′ region in addition to defects elsewhere in genes that are dispensable for growth in fibroblasts (11–18). The U1′/b′ region encodes a viral CXCL chemokine (gene UL146; ref. 19), a tumor necrosis factor receptor homolog (gene UL144; ref. 20), natural killer cell evasion functions (genes UL141 and UL142; refs. 21–23), a regulator of latency (gene UL138; refs. 24, 25), and several other uncharacterized functions. Therefore, the U1′/b′ region is likely to contribute to virulence via several pathways. Spontaneous defects clearly arise elsewhere in the HCMV genome, yet their significance is only occasionally defined functionally (13, 18, 26, 27). Indeed, whole-genome sequencing of multiple clinical strains following growth in vitro reveals that clinical HCMV genomes invariably mutate whether passed in fibroblasts, epithelial cells, or endothelial cells (28).

There is a need to develop and characterize low-passage HCMV strains for experimental applications. However, the generation of laboratory strains of WT HCMV is problematic not only because clinical samples often contain multiple strains (27, 29–36), but especially because genetic adaptations occur even during the early stages of passage in cell culture. Thus, most strains passaged in fibroblasts are mutated in 1 of 3 adjacent genes (UL128, UL130, and UL131A; collectively termed the UL128 locus, UL128L) whose products form a complex bound to glycoproteins gH and gL in the virion envelope (37–43). These mutations inhibit formation of the complex and thereby render the virus incapable of infecting epithelial, endothelial, and certain myeloid cell types (12, 44–48).
Figure 1
Steps in the construction of pAL1111. The approximate location of the insertion site of the BAC vector between US28 and US29 in the US region of the HCMV genome is shown at the top. The designations of BACs (pAL series) and viruses (RCMV series) are shown on the right. Boxes indicate protein-coding regions (labeled), and circles denote loxP sites (L). The BAC box represents pBeloBAC11, the eGFP/Puro box represents a cassette expressing eGFP and a puromycin resistance protein, and the Cre box represents a Cre recombinase gene containing a synthetic intron.
The genomes of several HCMV strains have been captured as BACs, including a number based on low-passage viruses, and these are proving to be an invaluable resource for research (49–54). However, none contains the full complement of HCMV genes, both because of mutations that occurred in cell culture prior to being cloned and, in most cases, because sequences were deleted in order to accommodate the BAC vector. In addition, the original clinical material is not available for most of these strains, thus preventing the sequences of the cloned genomes being verified against those of unpassaged virus. To provide a reliable source of HCMV gene sequences and a reproducible source of genetically intact, clonal virus for pathogenesis studies, we sought to produce an infectious BAC containing the complete HCMV genome. Cloning the HCMV strain Merlin genome led to the identification of disabling mutations in genes RL13 and UL128, which were repaired first singly and then in combination leading to the identification of disabling mutations in genes RL13 and UL128, which were repaired first singly and then in combination between virus genes US28 and US29. Preliminary studies (data not shown) indicated that straightforward insertion of the BAC-targeting vector into the HCMV genome by homologous recombination during virus growth in primary human fetal foreskin fibroblasts (HFFFs) was associated with compensating deletions in virus sequences, presumably due to genome size constraints operating during virus DNA packaging. The BAC-targeting vector was therefore redesigned to replace the region of the HCMV genome containing US29–US34 (Figure 1). BAC vector DNA was transfected into HFFFs, which were superinfected with HCMV strain Merlin (p5), and puromycin selection was used to enrich for recombinants. Circular DNA was extracted and electroporated into E. coli. A total of 22 clones were analyzed by restriction endonuclease digestion (data not shown), and all contained the HCMV genome minus US29–US34. Since both U₃ and U₄ may be present in either orientation in an HCMV genome, a linear molecule will be one of 4 isomers and a circular molecule one of 2. Restriction endonuclease profiles corresponding to both circular conformations were detected among the BAC clones (data not shown). A single clone (pAL1031; Figure 1) of the 12 clones with U₃ and U₄ present in the standard arrangement was selected for further analysis.

Recombineering (55, 56) was used to insert the US29–US34 region into pAL1031, thus generating pAL1040, which contains the entire Merlin genome (Figure 1). Cre/lox technology has been deployed previously to promote excision of prokaryotic vector sequences (also located between US28 and US29) from an Ad169 BAC following transfection into HFFFs (53). We adopted this strategy to produce a self-excising Merlin BAC by 2 rounds of recombineering, in which the enhanced GFP (eGFP)/Puro cassette in pAL1040 was replaced by a gene encoding Cre recombinase, thus generating pAL1053 (Figure 1). The version of Cre used contained a synthetic intron to prevent its expression in E. coli (57, 58). Following transfection into HFFFs however, Cre recombinase was expressed and mediated removal of the BAC vector by recombination between loxP sites engineered at the junctions with the virus genome. Thus, the only exogenous sequences remaining in virus generated from pAL1053 and subsequent BACs were those of single loxP and Nhel sites (40 bp in total) located between US28 and US29 (Figure 1).

Sequence of BAC pAL1053. Sequencing of the entire HCMV component of pAL1053 confirmed that the whole Merlin genome had been captured and that a single a/a′ sequence was present between

---

### Table 1

Sequence differences between Merlin (NC_006273) and the HCMV genome in pAL1053

<table>
<thead>
<tr>
<th>Position</th>
<th>Change (NC_006273 → pAL1053)</th>
<th>Coding region affected</th>
<th>Coding effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>11363</td>
<td>CAAAAAAC → CAAAAAAAC</td>
<td>RL13</td>
<td>f</td>
</tr>
<tr>
<td>40231</td>
<td>TTGGAGG → TTGGAGG</td>
<td>UL32</td>
<td>None</td>
</tr>
<tr>
<td>48828</td>
<td>GGCAGAC → GGCAAAC</td>
<td>UL36</td>
<td>R → C</td>
</tr>
<tr>
<td>49054</td>
<td>CATGATC → CATGATC</td>
<td>UL36</td>
<td>I → M</td>
</tr>
<tr>
<td>49203</td>
<td>GACAGTA → GACAGTA</td>
<td>UL36</td>
<td>R → C</td>
</tr>
<tr>
<td>49791</td>
<td>AGGGTG → AGGATG</td>
<td>UL36</td>
<td>P → S</td>
</tr>
<tr>
<td>49792</td>
<td>GGGTGAC → GGGTGTC</td>
<td>UL36</td>
<td>H → Q</td>
</tr>
<tr>
<td>194592–194643 (1025–1076)</td>
<td>ACACCCCGTCCCGGCACACCCCCCGTCCCGGCACACCCCCCGTCCCGG → G</td>
<td>None (b/b)</td>
<td>None</td>
</tr>
<tr>
<td>195063 (605)</td>
<td>GACGCAGC → GGAAGGC</td>
<td>None (b/b)</td>
<td>None</td>
</tr>
<tr>
<td>194781</td>
<td>CCGTAGA → CCGAGA</td>
<td>None (b only)</td>
<td>None</td>
</tr>
</tbody>
</table>

Nt positions are given with respect to NC_006273. f, frameshift. Altered bases underlined.
the b and c, and b′ and c′, sequences. As in Merlin (p3), pAL1053 contained a G to A (G → A) substitution at nt 176311 that introduces an in-frame stop codon and would lead to premature termination of the UL128 protein. pAL1053 differed from Merlin (p3) at a total of 12 locations (Table 1). These included 5 substitutions clustered in UL36. The sequences of PCR products amplified from the original clinical material, Merlin (p3), the earliest BAC precursor of pAL1053 (pAL1031), and 10 additional Merlin BACs from the same stage as pAL1031 showed that these mutations were present only in pAL1031, whereas the sequences from all other sources were identical to that of the Merlin (p3) sequence. Thus, the differences in UL36 were atypical and limited to the Merlin genome that was captured in pAL1031. The UL36 mutations in pAL1053 were repaired by recombineering to match the Merlin (p3) sequence, thus generating pAL1111 (Figure 1).

The most significant additional difference between Merlin (p3) and pAL1053 (and hence pAL1111) was a frameshift in RL13 at nt 11363 caused by a single nt insertion (Table 1; also see below). The remaining differences had no apparent effect on protein-coding potential. A synonymous substitution in gene UL32 was found retrospectively to represent a single nt polymorphism in the Merlin (p3) genome. Three alterations were noted in the b/b′ inverted repeat sequence. Two of these were substitutions, one being present in both b and b′ and the other in b′ alone. The other was a 51-bp deletion due to natural length variation in a tandem repeat sequence in both b and b′. The b/b′ sequence is the region of the HCMV genome that is most prone to variation in HCMV (18, 59)

Mutations in RL13. To investigate the origin of the frameshift mutation in the RL13 coding region (nt 11189–12070), this locus was sequenced in 10 additional BACs from the same stage as pAL1031. Surprisingly, all were found to contain disruptive mutations, and 4 different classes (classes 1–4) of mutant were identified, each of which is predicted to express a truncated RL13 protein (Figure 2). To determine whether RL13 was mutated in the virus stock used to generate pAL1031, the coding region was amplified by PCR from Merlin (p5) and 15 clones were sequenced (Figure 2). Mutations that were detected in single clones, except where they corresponded to mutations in the BACs, could have arisen by PCR error and were excluded from the analysis. The PCR clones identified mutations that corresponded to 3 of the 4 mutant groups
were used for Merlin (p5). Nine clones were WT in sequence, and the WT RL13 coding region. Merlin sequence was determined from bacteriophage M13 clones of the linear RCMV1111 genome. Infections of HFFFs at low MOI from Merlin at the pAL1031 stage were mutated in RL13 raised the US28 and US29. Intracellular FACS indicated that RCMV1111 expressed the US28 protein during the course of productive infection at an abundance comparable with Merlin (p5) (Figure 1). Since an antibody was not available for the US29 protein, US29 transcript levels expressed by RCMV1111 and Merlin (p5) were measured by quantitative RT-PCR (QRT-PCR) (Figure 3D), using UL123 (IE1) as a standard, and were comparable. Thus, the residual 40 bp in RCMV1111 had no discernible effect on expression of the adjacent US28 and US29 genes.

Expression of eGFP by virus generated from the Merlin BAC. To facilitate the characterization of viruses generated from Merlin BACs, an IRES-eGFP expression cassette was inserted into pAL1111 immediately downstream from UL122 (IE2), so that expression was under the control of the major IE promoter (60), generating pAL1158. In the resulting virus (RCMV1158), eGFP was expressed with IE kinetics (data not shown). In growth studies performed at low MOI (0.01 PFU/cell) in HFFFs, RCMV1158 exhibited a small but consistent reduction in levels of virus production (2-fold) at 12 days post-infection (PI) (Figure 3E).

Repair of RL13 or UL128 results in growth defects in fibroblasts. In order to reconstitute the complete HCMV gene complement, the lesions in UL128 and RL13 in pAL1111 were repaired seamlessly by recombineering, both singly and in combination, culminating in a set of 4 BACs (pAL1111, pAL1159, pAL1160, and pAL1128; Table 3). The complete sequence of the fully repaired BAC (pAL1128) was verified by Solexa sequencing. A corresponding set of eGFP-tagged BACs was also constructed (pAL1158, pAL1159, pAL1160, and pAL1161; Table 3).

To assess the efficiency of virus spread from cell to cell, the 4 eGFP-tagged BACs were transfected individually into HFFFs, the cells were overlaid, and then plaque areas were measured at 3 weeks post-transfection (PT) (Figure 4 and Figure 5A). Comparable results were obtained using the BACs lacking the eGFP tag (data not shown). Viruses with either RL13 (RCMV1159) or UL128 (RCMV1160) repaired produced much smaller plaques than the double mutant (RCMV1158). Furthermore, the virus (RCMV1161) with both genes repaired produced even smaller plaques. RL13 and UL128 thus act independently to restrict either cell-to-cell transmission or virus production in fibroblast cultures.

To investigate whether the mutants detected in Merlin (p3 and p5) were present in the clinical sample or whether they originated during cell culture, 10 PCR clones of RL13 were generated from the original clinical sample, using the same primers and conditions as were used for Merlin (p5). Nine clones were WT in sequence, and one clone contained a single synonymous substitution (A → G at nt 11959). This substitution was not detected in any versions of Merlin from passage in cell culture and thus most probably resulted from a PCR error. These data indicated that RL13 was predominately WT in the clinical sample and are consistent with RL13 mutants having arisen and been selected during cell culture.

<table>
<thead>
<tr>
<th>Position</th>
<th>Mutation (NC_006273 → clone)</th>
<th>Residue</th>
<th>Coding effect</th>
<th>No. of M13 clones</th>
<th>No. of PCR clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>11191</td>
<td>CATGGCAC → CATGGCAC</td>
<td>1</td>
<td>M → I</td>
<td>1/6</td>
<td>0/22</td>
</tr>
<tr>
<td>11537</td>
<td>GAAATGA → GAAATGA</td>
<td>117</td>
<td>V → M</td>
<td>4/8</td>
<td>11/22</td>
</tr>
<tr>
<td>11563</td>
<td>ATGGAAT → ATGCAAT</td>
<td>125</td>
<td>W → C</td>
<td>2/8</td>
<td>1/22</td>
</tr>
<tr>
<td>11563</td>
<td>ATGGAAAT → ATGCAAT</td>
<td>125</td>
<td>W → C</td>
<td>2/8</td>
<td>4/22</td>
</tr>
<tr>
<td>11900</td>
<td>ATGGAAA → ATGCAA</td>
<td>238</td>
<td>E → Q</td>
<td>2/8</td>
<td>4/22</td>
</tr>
<tr>
<td>11884</td>
<td>ATGCGGG → ATTTGGG</td>
<td>232</td>
<td>None</td>
<td>1/8</td>
<td>2/22</td>
</tr>
<tr>
<td>11964</td>
<td>TATGGGT → TATGGGT</td>
<td>259</td>
<td>S → X</td>
<td>1/8</td>
<td>2/22</td>
</tr>
</tbody>
</table>

Nt positions are given with respect to NC_006273. X, termination codon. Altered bases underlined.
relative plaque sizes (Figure 5A), with slowest growth occurring when RL13 and UL128 were both repaired (Figure 5, A–C). Repairing either RL13 or UL128 also gave a clear reduction in the levels of infectious virus released into the supernatant (Figure 5C). Repairing both genes delayed dissemination of the infection through the monolayer, but substantial and increasing amounts of infectious virus were released by 5–7 weeks PT (Figure 5C).

In titrations conducted to generate Figure 5C, the plaque sizes were largely in accord with those reported in Figure 5A. However, some plaques from the repaired viruses were substantially larger than the majority. Plaque sizes were measured at the earliest time point possible: 2 weeks PT for RCMV1158, 3 weeks PT for RCMV1159, and RCMV1160, and 4 weeks PT for RCMV1161 (Figure 5D). Two clearly distinct plaque sizes were evident for RCMV1159. The 3 larger plaques were expanded individually in HFFFs, and RL13 was amplified by PCR and sequenced. Each virus contained the same 2 mutations in RL13, specifically C → A at nt 11286 (resulting in S → Y) and C → G at nt 11436 (resulting in the introduction of an in-frame stop codon). Sequencing of viruses derived from the 3 largest RCMV1161 plaques did not reveal any mutations in RL13 or UL128.

Repairing RL13 results in growth defects in epithelial cells. Restoration of RL13 independently of UL128 suppressed virus replication in fibroblasts, with approximately 10% of virus having remutated in

Figure 3
Properties of Merlin BAC–derived virus RCMV1111. (A) Restriction endonuclease profiles of DNA extracted from Merlin (p5), RCMV1111 (the virus derived from pAL1111), and pAL1111. Markers (kb) are shown. (B) Virus titers released into the medium following infection of HFFFs at an MOI of 0.01 PFU/cell. Error bars show mean ± SD; results are representative of 2 experiments. (C) Intracellular FACS staining for the US28 protein in cells at various times PI with either Merlin (p5) or RCMV1111. Controls in the form of samples stained with IgG are shown. (D) Copy numbers of US29 transcripts relative to UL123 (IE1) transcripts at various times PI with Merlin (p5) or RCMV1111. Figures were averaged from 3 experiments; error bars show mean ± SD. (E) Virus titers released into the medium following infection of HFFFs at an MOI of 0.01 PFU/ml with RCMV1158 or RCMV1111. Error bars show mean ± SD; results are representative of 2 experiments.
RL13 by 3 weeks PT (see above; Figure 5D). To determine whether this effect was limited to fibroblasts or whether RL13 may be advantageous in other cell types (as is the case for UL128L), retinal pigmented epithelial cells (ARPE-19s) were transfected with pAL1160 and pAL1161 (Table 3) and the size of plaques was measured (Figure 6C). The lack of growth of UL128 mutants (RCMV1158 and RCMV1159) in ARPE-19s (data not shown) was consistent with previous observations that UL128L is essential for epithelial cell tropism. As was the case with HFFFs, plaques generated by viruses in which UL128 was repaired (RCMV1160 and RCMV1161) were eventually overgrown by uninfected cells, and virus dissemination required periodic reseeding of the monolayer. Both RCMV1160 and RCMV1161 grew much more slowly in ARPE-19s than HFFFs, and, unlike HFFFs, ARPE-19s did not attain complete infection (Figure 6B), instead reaching a plateau and forming a chronic infection. The virus in which both UL128 and RL13 were repaired (RCMV1161) grew much more slowly than RCMV1160 and formed a chronic infection in which a much lower proportion of cells was infected (approximately 15% for RCMV1161 compared with 35% for RCMV1160). RCMV1161 also produced lower levels of cell-free virus than RCMV1160 (Figure 6C). This effect was only partially attributable to the lower numbers of infected cells, since RCMV1161 infected about half the number of cells as did RCMV1160 and yet produced over 10-fold less cell-free virus. Comparable results were obtained using viruses lacking the eGFP marker. Thus, repairing RL13 results in slower cell-to-cell spread and reduced levels of cell-free virus from epithelial cells as well as fibroblasts.

**Rapid generation and selection of RL13 and UL128L mutants in fibroblasts.** To generate working stocks of BAC-derived viruses in HFFFs, the cells in a 25-cm² flask were transfected with BAC DNA and reseeded periodically until 100% were infected. Cell-free virus was transferred to a single 150 cm² flask and then to five 150 cm² flasks, and virus stocks (p3) were prepared. To test the genetic integrity of the virus, 10 PCR clones each of RL13 and UL128L were sequenced. Differences from the original sequences were excluded as probable PCR errors if they were detected in single clones. An analysis of 13 virus stocks (Table 4), each derived independently from BAC DNA, found that (a) when only RL13 was repaired, it mutated in all stocks (stocks 1–4); (b) when only UL128 was repaired, the UL128 locus mutated in 2 of 3 HFFF derived stocks (stocks 5, 6), whereas it remained intact in epithelial derived stocks (stocks 8, 9; a synonymous substitution was detected in one instance); and (c) when both RL13 and UL128 were repaired, RL13 mutated in 1 stock (stock 11), whereas the UL128 locus appeared to remain intact. When the stock (stock 11) containing an RL13+ UL128+ virus that harbored a frameshift in RL13 in 20% of genomes was passaged a fourth time, the RL13 mutation was then observed in 100% of clones, along with a deletion compromising both UL128 and UL130 in the UL128 locus. Thus, both RL13 and the UL128 locus remutated in fibroblasts, with the former tending to mutate more rapidly.

**RL13 encodes a virion envelope protein.** Since RL13 has not previously been characterized, it was important to establish and investigate its expression. RL13 was therefore tagged with a sequence encoding a C-terminal V5 epitope and inserted into a recombinant adeno virus (RAd) vector. The V5 epitope was also fused to RL13 within the C-terminal V5 epitope and inserted into a recombinant adenovirus (RAd) vector. The V5 epitope was also fused to RL13 within the

### Table 3

<table>
<thead>
<tr>
<th>BAC</th>
<th>Derived virus</th>
<th>RL13</th>
<th>UL128</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAL1111</td>
<td>RCMV1111</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pAL1119</td>
<td>RCMV1119</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pAL1120</td>
<td>RCMV1120</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>pAL1128</td>
<td>RCMV1128</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**BACs containing no tags**

<table>
<thead>
<tr>
<th>BAC</th>
<th>Derived virus</th>
<th>RL13</th>
<th>UL128</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAL1158</td>
<td>RCMV1158</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pAL1159</td>
<td>RCMV1159</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pAL1160</td>
<td>RCMV1160</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>pAL1161</td>
<td>RCMV1161</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**BACs tagged with eGFP**

<table>
<thead>
<tr>
<th>BAC</th>
<th>Derived virus</th>
<th>RL13</th>
<th>UL128</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAL1279</td>
<td>RCMV1279</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pAL1280</td>
<td>RCMV1280</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**BACs containing RL13 tagged with a V5 epitope**

<table>
<thead>
<tr>
<th>BAC</th>
<th>Derived virus</th>
<th>RL13</th>
<th>UL128</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAL1393</td>
<td>RCMV1393</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pAL1448</td>
<td>RCMV1448</td>
<td>(+et0)</td>
<td>–</td>
</tr>
<tr>
<td>pAL1498</td>
<td>RCMV1498</td>
<td>(+et0)</td>
<td>(+et0)</td>
</tr>
</tbody>
</table>

**BACs tagged with eGFP, tet Operators before RL13/UL128L**

<table>
<thead>
<tr>
<th>BAC</th>
<th>Derived virus</th>
<th>RL13</th>
<th>UL128</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAL1158</td>
<td>RCMV1158</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pAL1159</td>
<td>RCMV1159</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pAL1160</td>
<td>RCMV1160</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>pAL1161</td>
<td>RCMV1161</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**Figure 4**

**Growth of BAC-derived viruses in HFFFs.** Images show eGFP expressed by plaques formed in HFFFs 3 weeks PT with BAC DNA as indicated. Scale bars: 100 μm.
tant to EndoH digestion and are thus presumably fully mature. PNGaseF digestion reduced the 55-kDa proteins to 35 kDa and, in addition, reduced the 80- and 100-kDa proteins to 57 and 65 kDa, respectively. Since neither enzyme was able to reduce the 80- and 100-kDa proteins to 35 kDa, it is likely that mature gpRL13 contains O-linked, as well as N-linked, sugars. Identical results were obtained with cells infected with RCMV1280 (RL13V5+UL128+), and no bands were seen using lysates from RL13– virus (data not shown). The observation that gpRL13 appears to be more extensively glycosylated in the context of HCMV infection than when expressed in isolation is consistent with previous observations on CD155 and gpUL18, to the effect that HCMV infection alters normal cellular glycosylation processes (22, 61).

The V5 epitope was also used to track gpRL13 expression by immunofluorescence. When RL13V5 was expressed using the RAd vector, gpRL13 trafficked to discrete cytoplasmic vesicles, a proportion of which could be stained with the early endosomal marker Rab5A (Figure 7C). This observation is consistent with much of the protein having transited the Golgi apparatus. In the context of a productive HCMV infection, the intracellular distribution of gpRL13 was different. The protein localized to the cytoplasmic site of virion assembly and colocalized in both HFFFs and ARPE-19s with a marker for the trans-Golgi network (TGN46) and pp28 (an outer tegument protein encoded by gene UL99 that interacts with the envelope and is acquired by the virion in the cytoplasm; Figure 7D) and gH (a virion envelope glycoprotein encoded by gene UL75; Figure 7E).

The finding that gpRL13 is a glycoprotein that tracks to the site of virion assembly raised the possibility that it might be a virion surface envelope protein. Growth of RL13+ HCMV in vitro is inefficient, and spontaneous mutants will arise and be rapidly selected (see above). Nevertheless, we found that limited, short-term growth of RL13+ virus could be fostered by sequential reseeding of cultures (Figure 5), and the antibody to a C-terminal tag would preferentially recognize full-length, nonmutated gpRL13 even if a proportion of mutants had arisen during passage. In virions purified from such cultures, only the 100-kDa mature, Endo H–resistant form of gpRL13 was detected. In virion fractionation studies, this protein copurified with glycoprotein B (gB; the product of gene UL55) in the soluble envelope fraction, rather than with pp65 (the product of gene UL83) in the tegument (Figure 8A). The V5 epitope is predicted to be located topologically on the inner side of the envelope, and this orientation was supported by immunoelectron microscopy, in which gold-labeled secondary antibody exhibited extensive labeling around the inner surface of the envelope (Figure 8B).

Conditional expression of RL13 and UL128L. The BAC pAL1111 (RL13–UL128+) has utility as a reproducible source of fully characterized, clonal HCMV that exhibits good genetic stability and rapid growth to high titers in fibroblasts. We also show above that short-term experiments can be performed using the Merlin BAC in which RL13 and UL128 are repaired. The major phenotypic impact these genes have on HCMV biology emphasized the need to work with WT virus. However, the rapid emergence of mutants at these loci during amplification from Merlin BAC
transfections meant virus stocks were inevitably contaminated with mutants. In order to enable studies with phenotypically WT virus, we sought to suppress expression of RL13 and the UL128 locus during virus expansion following Merlin BAC transfection. To achieve this, hTERT-immortalized HFFFs were transduced with a retrovirus encoding the tet repressor containing a C-terminal nuclear localization signal (NLS) and tet operators were inserted upstream of the translation initiation codon of the gene to be modulated (62, 63). In HFFF-tet cells, the tet repressor binds to the tet operators and prevents transcription of the target gene, whereas transcription proceeds as normal in parental HFFFs. When tested with a RAd-expressing tet-regulated GFP, expression levels in HFFF-tet cells were reduced 180-fold relative to parental HFFFs, showing little more fluorescence than cells infected with an empty vector control (Figure 9A).

The transcriptional initiation sites for RL13 are located approximately 30- and 110-bp upstream from the translational initiation codon (K. Baluchova, unpublished observations). Transcription of UL128 starts within the UL130 coding region, and an mRNA initiating approximately 10 bp upstream of UL131A potentially encodes both the UL131A and UL130 proteins (26, 43). Tet operators were inserted in various locations upstream from the RL13 or UL131A coding regions, and their capacity to selectively promote replication of RL13’ or UL128’ viruses in HFFF-tet cells was tested empirically. The optimal location for the tet operator was determined to be a single tet operator 19 bp upstream of RL13 (pAL1448) and 2 tet operators 33 bp upstream of UL131A (pAL1393; Table 3). Plaque formation was inhibited approximately 10-fold in HFFF cells by the presence of an intact UL128 locus, and this inhibition was relieved by infection of HFF-tet cells with a tet-controlled UL131A virus, RCMV1393 (Figure 9, B and C). Similarly, plaque formation was inhibited approximately 10-fold in HFFF cells by the presence of an intact RL13, and this inhibition was relieved by infection of HFFF-tet cells with a tet-controlled RL13 virus, RCMV1448 (Figure 9, D and E). Titers of tet-controlled viruses obtained from HFFF-tet cells were increased 18- to 32-fold relative to the non-tet–controlled viruses (Figure 9F). In addition, sequencing of 10 PCR clones of UL128L and RL13 showed all contained the WT sequence. Following passage in HFFs with this RCMV1393 stock, virus was able to infect ARPE19 cells and produce plaques of a size equivalent to the non–tet-controlled parental virus RCMV1160 (Figure 9G). Finally, having verified that both RL13 and UL131A could be independently tet-controlled, a virus was constructed in which both genes contained tet operators upstream of their ATG codons. Like the previous constructs, plaque formation was inhibited in HFFF cells by the presence of an intact UL128 and UL128 locus, and this repression was relieved by infection of HFFF-tet cells with a virus in which both RL13 and UL128 were tet-controlled (RCMV1498) (Figure 10, A and B).

Thus, repression of RL13 and UL131A in HFFF-tet cells demonstrably provides a means to amplify Merlin virions containing intact versions of RL13 and the UL128 locus. Two options are available to produce phenotypically WT HCMV from these viruses: either transcriptional repression can be released by the addition of doxycycline to infected HFFF-tet cells or, as demonstrated above, a final replication cycle can be performed in HFFs.

Discussion

Diagnostic laboratories have long recognized that HCMV strains in clinical samples do not adapt readily to cell culture. Our construction of what we believe is the first BAC containing a complete, characterized copy of a genetically intact HCMV genome provides an explanation for this phenomenon in showing that adaptation is dependent on 2 independent mutations, one in the UL128 locus (previously recognized; ref. 26) and one in RL13 (identifiable in the present study). To date, HCMV research has by necessity used viruses that, to varying degrees, are compromised genetically. Although laboratory-adapted viruses have proved invaluable, there is a clear need to develop systems that represent the WT virus responsible for clinical disease. To this end, the Merlin genome (p5) was cloned using BAC technology, and the sequence of an initial BAC (pAL1053) was compared with that of the parental strain. Specific issues relating to tissue culture adaptation were clarified and resolved using sequence data derived directly from the original clini-
derived from a small proportion of mutants that arose in cell cul
to pAL1053 and its predecessor, pAL1031, and were presumably
what we believe is the first time, enabled experiments to be under
HCMV strain Merlin (RCMV1111; RL13
pAL1161 contain the complete HCMV gene complement and, for
UL36, UL128, and RL13. The differences in UL36 were specific
viruses was associated with rapid selection of further mutants at
these loci. Therefore, we developed the HFF-tet cell line, which
allowed RL13 or UL131A to be repressed during recovery of virus
from the BAC, thus providing a stable platform for experimenta
tion on genetically intact HCMV.

Although HCMV replication has been described in a wide range
of cell types, efficient virus production in vitro has been achieved
only using viruses adapted to growth in fibroblasts. Indeed, exist-
ing cell-culture systems appear to be essentially incompatible
with efficient propagation of WT HCMV. HCMV is not excep-
tional in this regard. Of the 8 known human herpesviruses, only
WT herpes simplex virus types 1 and 2 (HSV-1 and HSV-2) can be
propagated readily to high titer in cell culture. Nonetheless, very
high HCMV virus loads are often detected in urine samples from
congenitally infected neonates (as was the case for the sample
from which Merlin was isolated), and this indicates that condi-
tions exist in vivo that are compatible with efficient replication of
WT HCMV. It is possible that RL13 regulates a switch to pro-
ductive HCMV replication in vivo.

The UL128 locus is essential for the efficient infection of ep-
ithelial and endothelial cells by HCMV (37–43) but is detrimental
to growth in fibroblasts. Consequently, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its
tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its
tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

tergent growth properties in fibroblasts. Consequence, mutants in this locus are
selected when clinical isolates are passaged in fibroblasts (12, 26,
27, 44, 45, 47, 48). Merlin was originally selected for detailed inves-
tigation from a series of clinical isolates partly on the basis of its

Table 4
RL13 and UL128L mutations detected by PCR in BAC-derived viruses at p3

<table>
<thead>
<tr>
<th>Stock</th>
<th>RL13 mutations</th>
<th>UL128L mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock</td>
<td>Mutations</td>
<td>Position</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>1</td>
<td>G → T</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>2</td>
<td>ΔA</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>3</td>
<td>G → A</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>4</td>
<td>A → G</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>5</td>
<td>C → T</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>6</td>
<td>T → C</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>7</td>
<td>CAAGA → TCTTG</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into ARPE-19s</td>
<td>8</td>
<td>ΔT</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>9</td>
<td>A → G</td>
</tr>
<tr>
<td>RL13-UL128- transfected into HFFFs</td>
<td>10</td>
<td>None</td>
</tr>
<tr>
<td>RL13-UL128- transfected into HFFFs</td>
<td>11</td>
<td>G → T</td>
</tr>
<tr>
<td>RL13-UL128- transfected into HFFFs</td>
<td>12</td>
<td>None</td>
</tr>
<tr>
<td>RL13-UL128- transfected into HFFFs</td>
<td>13</td>
<td>None</td>
</tr>
</tbody>
</table>

Each row represents an individual virus stock. Nt positions are given with respect to NC_006273. +, WT gene; –, mutated gene; ^, inserted nt; Δ, deleted nt.
each serial subculture (usually weekly) of an infected cell monolayer. However, in the present study, virus passage involved destruction of the entire cell monolayer following infection with cell-free virus, a process that normally involves many more cycles of virus replication. Also, given that viruses mutated in the UL128 locus released higher amounts of virus into the medium, our use of cell-free virus may have further encouraged the selection of mutants.

The identification of an RL13 mutation in pAL1053 was unexpected because the genome sequence of Merlin (p3) indicated that this coding region was intact (27). It led to the discovery that low-passage Merlin stocks actually consisted of mixtures of various different RL13 mutants. This finding brings a resolution with observations that almost all HCMV strains and BACs are overtly mutated in RL13 (27). Parallels exist between RL13 and the UL128 locus in the fact that mutations arise in both genes during adaptation to culture in fibroblasts. However, it is clear that RL13 and the UL128 locus can function independently because either gene in intact form was capable of suppressing virus replication, and

Figure 7
Characterization of gpRL13. (A and B) Western blot performed on RL13V5, showing the sizes of the proteins in native form or following digestion with EndoH or PNGaseF. RL13V5 was either expressed in isolation from a RAd (A) or from its native position in RCMV1279 (B). (C) Immunofluorescence performed for the indicated antigens on HFFFs infected with the RAd expressing RL13V5 (×640 magnification). (D and E) Immunofluorescence performed for the indicated antigens either on HFFFs infected with RCMV1279 or on ARPE-19s infected with RCMV1280. Original magnification, ×640.
repair of both had an additive effect. The RL13 phenotype was also shown to be distinct from that of the UL128 locus, since RL13 suppressed HCMV replication in both fibroblasts and epithelial cells, whereas the UL128 locus was not only stable, but promoted infection, in epithelial cells. Nevertheless, the finding that gpRL13 is present in the virion indicates that it may have a role in modifying tropism (similar to the UL128 locus) or in modulating cell signaling during virus entry (as proposed for gB).

The apparent ease with which RL13 mutants were selected in cell culture raised the possibility that they preexisted in the clinical sample, perhaps reflecting the potential for an expanded cell tropism in vivo. This possibility is technically challenging to disprove, since such mutants might be present in very low proportions. However, we were unable to detect the existence of RL13 mutants in the primary clinical sample from which Merlin was derived. It was estimated that approximately 10³ particles were used in the initial infection of a 25-cm² flask to isolate Merlin, and this would have resulted in most cells receiving a particle. Nonetheless, passage 1 of Merlin took 4 weeks to complete. The mutation rate of DNA viruses is relatively low and has been calculated at 0.003 per genome replication for HSV-1 (65). If this value holds for HCMV, mutants in the UL128 locus and/or RL13 would be expected to arise during p1. This mutation rate operating on 10⁶ viruses would result in a total of 3,000 mutations per genome replication. Since RL13 accounts for 879/235,646 bp in strain Merlin genome, approximately 11 mutations would be in RL13. The acquisition of mutations in both RL13 and UL128 could have resulted either from sequential mutations in the same template or as a result of recombination between independent mutants. Despite the counterintuitive nature of such rapid selection events in a herpesvirus, mutations in RL13 and the UL128 locus were also identified at p3 following transfection of the repaired BACs into HFFs, and these were clearly generated de novo. Interestingly, mutants emerged in RL13 more rapidly than in the UL128 locus. Some data also indicated that RL13 may mutate more rapidly in the context of UL128 driving its genetic variation remains operational.

In support of our conclusions concerning the instability of RL13 in cell culture, a collaborative study has shown that RL13 mutants were invariably selected in fibroblasts, epithelial, and endothelial cells during sequential passage of HCMV strains from clinical samples (28). Also, in most existing HCMV BACs (including FIX, Ph, Towne, Toledo, and AD169), RL13 is mutated overtly by deletions, frameshifts, or substitutions that introduce premature termination codons (27), and that derived from strain TB40/E (54) contains a unique substitution that greatly reduces the prediction of a signal peptide for gpRL13 (A.J. Davison, unpublished observations). Relating to this, we note that on 2 occasions, viruses derived from Merlin BACs replicated relatively efficiently in fibroblasts and yet retained an intact RL13 (RCMV1159 [data not shown] and a proportion of RCMV1161; see Figure 5D). This suggests that another HCMV gene may be required for RL13 function and that this gene is less prone to mutation than RL13.

HCMV exhibits the highest degree of intrastrain sequence variation of any human herpesvirus, and RL13 is a member of a small group of genes that exhibit the greatest variation (27). Also among this group are genes UL146 (27, 35, 36, 66–68) and UL74 (gO) (27, 35, 36, 69–71), which are well characterized because of their utility in genotyping clinical isolates. The selection pressures responsible for generating such degrees of divergence are not fully understood, but their origins, and perhaps the era in which they have operated, appears to be ancient (66). Despite remarkable sequence variation among HCMV strains, UL146 and UL74 are stable within individual patients, and identical UL146 sequences have been detected in geographically distant individuals (35, 66). RL13 is one of 14 members of the RL11 gene family, which is believed to have arisen through gene duplication and then diverged during the evolution of the primate cytomegaloviruses (72). Several other members of this family are also hypervariable. As a virion envelope protein, gpRL13 may be a prime target for neutralizing antibody, and it is possible that selection is exerted on it in vivo to drive escape from the humoral immune response. In support of this, when CD4⁺ T cell responses to HCMV were measured, RL13 was one of the top 5 most immunogenic HCMV ORFs (73). It would be interesting to determine whether RL13 is now stable or whether the mechanism driving its genetic variation remains operational.

The question arises of why RL13 is detrimental to virus replication in both epithelial and fibroblast cells in culture. The virion components encoded by the UL128 locus are clearly also detrimental in fibroblasts, yet they are essential for infection of other cell types. Since gpRL13 is also present in the virion, it may likewise modify tropism in some way that has not yet been recapitulated in vitro. Recent studies have suggested that HCMV is capable in vivo of establishing persistent, low-level infections that are cor-
Figure 9
Repression of UL128L and RL13 by BAC-derived viruses in HFFF-tet cells. (A) FACS analysis of parental HFFFs or HFFF-tet cells, infected with empty control adenovirus (RAd-Ctrl) or RAd-expressing eGFP (RAd-GFP). RAd-GFP expresses GFP from the HCMV MIE promoter, containing 2 tet operators 10 bp downstream of the TATA box. (B–E) Plaque sizes of viruses generated at 2 weeks PT in HFFFs (B and D) or HFFF-tet cells (C and E), with cells overlaid to prevent cell-free spread of virus, showing (B and C) repression of UL128L in RCMV1393 and (D and E) repression of RL13 in RCMV1448. (F) Titers of virus stocks obtained from HFFF-tet cells infected with viruses in which RL13 and UL131A were tet controlled (RCMV1448 and RCMV1393, respectively) or the parental viruses in which RL13 and UL131A were not tet controlled (RCMV1159 and RCMV1160, respectively). (G) Plaque size of parental virus (RCMV1160) or virus in which UL131A was tet controlled (RCMV1393) in ARPE19 cells, 3 weeks PT.
related with the presence of glioblastomas (74–76). Alternatively, therefore, RL13 may act as a regulator, promoting persistence of HCMV by suppressing the switch to full lytic infection until virus dissemination is required. Clearly there are situations in vivo, as illustrated by UL138 (which promotes latency by suppressing lytic infection in CD34⁺ myeloid progenitors), where it can be beneficial for the virus to restrict or limit productive infection.

Consideration has been given to using live HCMV as a vaccination agent, vaccine carrier, and vector for gene therapy. Such studies are being developed with strains adapted to cell culture and have lost RL13 function. When the low passage HCMV strain Merlin (p5) (27) was utilized to generate BACs as described previously (62). Maxipreps were produced using NucleoBond BAC100 kits (Machery Nagel) from 500 ml overnight bacterial cultures. BAC DNA transfections (2 μg) of HFFs (10⁶ cells) were performed using a basic fibroblast kit and a Nucleofector (Amaxa) according to the manufacturer’s protocol. Transfections into ARPE-19s were performed in 25-cm² flasks using effectene (QIAGEN) according to the manufacturer’s protocol for 60-mm dishes.

Transformation of E. coli was performed by electroporation at 2.5 kV using a basic fibroblast kit and a Micropulser (Bio-Rad). When PCR products were cloned for sequencing, DNA was purified from agarose gels using a GFX PCR DNA and Gel Band Purification kit (GE Healthcare) and cloned into the manufacturer’s protocol for 60-mm dishes.

**Methods**

All studies in human and animal cells were approved by Bro Taf Local Research Ethics Committee, Cardiff, United Kingdom.

**Cells and viruses.** HFFs and ARPE-19s (77) were grown in DMEM ( Gibco; Invitrogen) containing 10% FCS (Gibco; Invitrogen) at 37°C in 5% CO₂. HCMV strain Merlin (p5) (27) was utilized to generate BACs as described below. Infections were performed as described previously (78), and viruses were titrated in triplicate by plaque assay for 14 days on HFFs using a 1% antibiotic (10⁶ cells) and sizes were computed using Openlab 3 software (Improvision).

**Construction of the BAC-targeting vector.** The BAC-targeting vector (pAL1026; Figure 1) used to capture the Merlin genome contained markers for eGFP and puromycin resistance (eGFP/Puro) and homology arms containing HCMV genes US28 and US34A (817 and 888 bp, respectively), a chloramphenicol resistance gene (cat), genes sopA, sopB, and sopC, which ensure active partitioning during cell division, and gene repE, which mediates assembly of the replication complex in E. coli at Ori2. Each homology arm was flanked by a loxP site, and a unique Nhel restriction site was positioned between them.

To generate pAL1026, pBetolBAC11 (New England Biolabs; GenBank U51113; ref. 80) was digested with Hpal and ApaLI, blunted-ended with Klenow DNA polymerase, and religated to remove the existing cos and loxP sites, thus generating pAL767. Homology arms matching US28 and US34 were amplified from Merlin DNA using primers US28F (GGCC-GCTAGCTGGCGACGTCGGATTCAATG; NheI site underlined), US28R (GGCC-GGATCCATACTCGTATATGATCTATACGAGTTATACGGTATAAT; BamHI site underlined), and US34F (GGCC-GGATCCATACTCGTATATGATCTATACGAGTTATACGGTATAAT; BamHI site underlined) and inserted into pAL767, thus generating pAL799. A promoter and linked by an internal ribosome entry site [IRES]) was inserted into the BAC-targeting vector YD-C19 (duplex HindIII site, and a unique Nhel restriction site was positioned between them.

**PCR.** Three DNA polymerases were used in PCR reactions according to the manufacturers’ protocols: Phusion (NEB) for fragments greater than 4 kb; Advantage 2 (Clontech) for amplification directly from virus stocks or cultures; and Expand Hi-Fi (Roche) for all other experiments. Oligonucleotide primers were purchased from Sigma-Aldrich at desalted purity. RL13 and the UL128 locus were amplified for sequencing using primers RL13F (ATCCTGAACATGAAGACTGACGTT) and RL13R (GAATTTACCATGATACAGTCC). Three DNA polymerases were used in PCR reactions according to the manufacturer’s protocol: Phusion (NEB) for fragments greater than 300 bp, Advantage 2 (Clontech) for amplification directly from virus stocks or cultures; and Expand Hi-Fi (Roche) for all other experiments. Oligonucleotide primers were purchased from Sigma-Aldrich at desalted purity. RL13 and the UL128 locus were amplified for sequencing using primers RL13F and RL13R, respectively.

**Isolation, transformation, and transfection of DNA.** Minipreps of plasmids or BACs were produced as described previously (62). Maxipreps were produced using NucleoBond BAC100 kits (Machery Nagel) from 500 ml overnight bacterial cultures. BAC DNA transfections (2 μg) of HFFs (10⁶ cells) were performed using a basic fibroblast kit and a Nucleofector (Amaxa) according to the manufacturer’s protocol. Transfections into ARPE-19s were performed in 25-cm² flasks using effectene (QIAGEN) according to the manufacturer’s protocol for 60-mm dishes.

**Construction of the BAC-targeting vector.** The BAC-targeting vector (pAL1026; Figure 1) used to capture the Merlin genome contained markers for eGFP and puromycin resistance (eGFP/Puro) and homology arms matching HCMV genes US28 and US34A (817 and 888 bp, respectively), a chloramphenicol resistance gene (cat), genes sopA, sopB, and sopC, which ensure active partitioning during cell division, and gene repE, which mediates assembly of the replication complex in E. coli at Ori2. Each homology arm was flanked by a loxP site, and a unique Nhel restriction site was positioned between them.

To generate pAL1026, pBetolBAC11 (New England Biolabs; GenBank U51113; ref. 80) was digested with Hpal and ApaLI, blunted-ended with Klenow DNA polymerase, and religated to remove the existing cos and loxP sites, thus generating pAL767. Homology arms matching US28 and US34 were amplified from Merlin DNA using primers US28F (GGCC-GCTAGCTGGCGACGTCGGATTCAATG; NheI site underlined), US28R (GGCC-GGATCCATACTCGTATATGATCTATACGAGTTATACGGTATAAT; BamHI site underlined) and US34F (GGCC-GGATCCATACTCGTATATGATCTATACGAGTTATACGGTATAAT; BamHI site underlined) and inserted into pAL767, thus generating pAL799. A promoter and linked by an internal ribosome entry site [IRES]) was inserted into the BAC-targeting vector YD-C19 (duplex HindIII site, and a unique Nhel restriction site was positioned between them.

**Graphs.** The Journal of Clinical Investigation (http://www.jci.org) Volume 120 Number 9 September 2010

![Graph A](http://example.com/graphA.png)

**Table 1**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RL13</td>
<td>Regulator, promoting persistence</td>
</tr>
<tr>
<td>UL128</td>
<td>Repressor, suppresses lytic</td>
</tr>
<tr>
<td>US28</td>
<td>Chloramphenicol resistance gene</td>
</tr>
<tr>
<td>US34</td>
<td>SopA, sopB, and sopC</td>
</tr>
</tbody>
</table>

**Figure 10**

Simultaneous repression of UL128L and UL13 by BAC-derived viruses in HFF-2 cells. Plaque sizes of viruses generated at 3 weeks PT in HFFs (A) or HFF-2 (B) cells, with cells overlaid to prevent cell-free spread of virus. Data points represent individual plaque sizes recorded for each mutant.
underlined). The amplified fragment was digested with PaeI and inserted into PacI-digested pAL815, thus generating pAL1026.

Construction of a Merlin BAC. 1 μg pAL1026 was linearized by NheI digestion and transected using Effectene (Qiagen) into 5 × 10^6 HEK293 cells. Cells were infected at 24 hours PT with Merlin (p5) at an MOI of 10, and recombinants were enriched by selection with puromycin (2.5 μg/mL) and detected by visualizing eGFP. When a significant proportion of cells exhibited eGFP expression, circular DNA was extracted using Hirt extraction (53, 81) and transfected into E. coli. Selection with chloramphenicol allowed the identification of BAC colonies, from among which pAL1031 (Figure 1) was analyzed in detail and repaired by recombineering.

Repairing of Merlin BACs. Recombineering was performed in E. coli SW102 using lacZ/amp/sacB (62) and galk selection cassettes (82) as previously described. The sequences of the primers used are listed in Supplemental Table 1 (supplemental material available online with this article; doi:10.1172/JCI42955DS1). All constructs were verified by sequencing modified regions and by restriction digest at each step. To enable the Merlin UL29–UL34 sequence to be reintroduced into pAL1031, the lacZ/amp/sacB cassette was amplified and inserted between the loxp site and US34A using primers SacBF-LoxPHom and SacBR-US34AHom to generate pAL1032 (Figure 1). To generate the insert, 2 PCR products were produced, digested with NheI, and ligated together. A 482-bp region encompassing the loxp site adjacent to US34 in pAL1026 was amplified using primers LoxP-R (GGCCGGTACGTAACCTCGTATTAATGTATGCTATACGAAATTTGAGATCGAGCTTGTATC; NheI site underlined) and Chlor-F (GGCCGAGCTCTGCGAAGATGCTTAAATG; BglII site underlined). A 5.2-kb region of the Merlin genome spanning genes US29–US28 was also amplified using primers US29F (GGCCGCTAGCCACCTCGGCAATTTCTTTCAA; NheI site underlined) and US34AR (GGCCGAGCTCTGGAGTTCTATATGGATATTG; BglII site underlined). Each amplicon was cloned into pCR2.1-TOPO (Invitrogen) for sequencing. A clone of each amplicon in which the sequence was correct was digested with BglII/NheI. The fragments were ligated together and recombineered into pAL1033, thus replacing the lacZ/amp/sacB cassette by US29–US34A and generating pAL1040 (Figure 1).

The eGFP/Puro cassette in pAL1040 was replaced by a Cre recombinease gene under the control of an SV40 promoter and containing a synthetic intron (57) so that the protein would be expressed in mammalian cells but not E. coli. The lacZ/amp/sacB cassette was amplified using primers SacBF-GFPProHom and SacBR-GFPProHom and inserted into pAL1040, thus generating pAL1047 (Figure 1). The Cre recombinease expression cassette was amplified from YD-C66 (a gift from D. Yu; ref. 53), using primers CreF-SV40Hom and CreR-BACHom, and recombineered in place of the lacZ/amp/sacB cassette, thus generating pAL1053 (Figure 1).

The sequence of pAL1053 was determined (see below) and indicated the presence of several differences from that of WT Merlin (see Results). Redundant sequences containing residual lacZ sequences were deleted by replacement with a galK selectable marker, which was amplified using primers GalKF-CreHom and GalKR-SopCHom. The galK marker was then removed using the primer RemoveGalKBAC, thus generating pAL1090 (Figure 1). To repair the lesions in gene UL36, the lacZ/amp/sacB cassette was amplified using primers SacBF-UL36Hom and SacBR-UL36Hom and inserted in place of UL36. UL36 was then amplified from Merlin DNA using primers UL36F and UL36R and recombineered in place of the lacZ/amp/sacB cassette, thus generating pAL1111 (Figure 1). Additional BACs were generated by repairing either and then both of the UL128 and RL13 mutations by the same technique. In each case, the lacZ/amp/sacB cassette was amplified using primers flanking the mutation (SacBF-UL128Hom and SacBR-UL128Hom or SacBF-RL13Hom and SacBR-RL13Hom) and recombineered into the relevant BAC, and the selectable marker was excised and the mutation repaired by recombination with oligonucleotide UL128Rep or RL13Rep.

To insert an IRES-eGFP expression cassette immediately downstream from gene UL122 (IE2) into each of the 4 BACs described above, the lacZ/amp/sacB cassette was amplified using primers SacBF-IE2Hom and SacBR-IE2Hom. In the resulting BACs, the lacZ/amp/sacB cassette was replaced by an IRES-eGFP cassette amplified from pIRES2-GFP (Clontech) using primers IRESF-IE2Hom and GFPRI-IE2Hom.

DNA sequencing. All PCR products in directly cloned or recombineered form were verified by sequencing using a BigDye 3.1 kit (ABI) and standard techniques, except that the number of cycles during sequencing was increased from 25 to 100. Reactions were purified using Dye Terminator Removal Columns (EdgeBio) and analyzed on an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems).

Two BACs were sequenced using published approaches (59). The complete HCMV component of pAL1053 was determined via standard Staden sequencing of a set of overlapping PCR products, and that of pAL128 in its entirety was determined using data from an Illumina Genome Analyzer (The GenePool, University of Edinburgh). The sequence of pAL1128 was deposited in GenBank as accession GU179001. All nts in the text are specified relative to Merlin (NC_006273).

QRT-PCR. QRT-PCR was performed to analyze the expression levels of gene US29 in infected cells. The cells were trypsinized and washed once in PBS, and whole-cell RNA was extracted using an RNeasy Mini Kit (Qiagen) and DNase-treated using Turbo DNA-free (Ambion). US29 transcripts were amplified using primers US29F (GACATCGTGGACAGCCCTCA) and US29R (CTTGGGCTTCCAGAGCCGC), and UL123 (IE1) transcripts were amplified using primers IE1F (GGGAAGAAGGTGAACAGAGTGA) and IE1R (TTCCCTACACCATCTCC). For RT-PCR with real-time detection, an iScript One-Step RT-PCR kit with SYBR Green (Bio-Rad) was employed according to manufacturer's instructions, and samples were analyzed on an iCycler (Bio-Rad). Serial dilutions of a Merlin BAC (pAL1053; see below) were used to establish a calibration curve. Since transcripts antisense to US29 have been reported (83), the sense transcript was amplified specifically using the reverse primers (100 pmol) with 100 ng RNA in a reverse transcription step at 50°C for 10 minutes, followed by inactivating reverse transcriptase at 95°C for 5 minutes. The forward primer was then added, and the samples were cycled 40 times through 95°C for 10 seconds, 60°C for 30 seconds, and 83°C for 10 seconds, with data being collected following the final step of each cycle. Melt curve analysis was performed following the 40 cycles, and amplification of a product of the correct size was confirmed by agarose gel electrophoresis. Samples were analyzed in triplicate, and the results were averaged. Samples from uninfected cells and infected cells processed without conducting the reverse transcriptase step were uniformly negative.

Flow cytometry. A Cytofix/Cytoperm plus kit (BD) was used for intracellular staining of the US28 protein, and the Tub-45 antibody (a gift from U. Höpken; ref. 84) was used (1:5) in combination with goat anti-mouse Alexa Fluor 647 (1:200) (Invitrogen). Samples were measured on a FACSCalibur (BD), and data were analyzed in FlowJo (Treestar). Generation of the HFF-tet cell line. The coding region of the tet repressor was amplified using primers tetR-F (GGGCGGATTCTATTGGCCGAATGATGCTCTA; BamHI site underlined, initiation codon in bold) and tetR-R (GGGCGGATTCTATTGGCCGAATGATGCTCTA; BamHI site underlined, initiation codon in bold) and cloned into the BamHI/EcoRI sites of the retrovirus vector pMXs-IP (85). Retrovirus stocks were produced by transfecting 293 Phoenix
Tagging RL13 in BACs with a V5 epitope. A sequence encoding a V5 epitope was fused to the C terminus of gpRL13 by recombineering. The lacZ/amp/sacB cassette was amplified with primers SacBF-RL13Hom and SacBR-RL13Hom and inserted into the BAC vector. After transformation of the E. coli strain DH10B containing the BAC, the lacZ/amp/sacB cassette was amplified with primers SacBF-UL131A and SacBR-UL131A and inserted into the BAC and then removed with oligonucleotide tetO-RL13, leaving behind 1 tet operator. To insert 2 tet operators 33 bp upstream of the UL131A coding region, the lacZ/amp/sacB cassette was amplified with primers SacBF-UL131A and SacBR-UL131A and inserted into the BAC and then removed with oligonucleotide tetO-UL131A, leaving behind 2 tet operators.

Generating replication-deficient RAd vectors. RL13 was cloned by recombineering into a RAd using the Ad2 system as described previously (62), using primers RL13SF and RL13SR, so that the gpRL13 was tagged with a C-terminal V5 epitope. Virus was recovered from the vector by transfection into 293TREx cells and titrated as described previously (89). Particle concentrations in the preparation were measured by laser illumination in a Zeiss LMS 510 confocal microscope, and images were captured using LSM software.

Acknowledgments

This work was supported by the Wellcome Trust, the United Kingdom Medical Research Council (MRC), and the Biotechnology and Biological Sciences Research Council (BBSRC). The authors thank Sian Llewellyn-Lacey and Carole Rickards for tissue culture support; Brent Ryckman for helpful discussions; Neal Copeland, Dong Yu, and Uta Hopken for reagents; Richard Darley and Phillip Taylor for reagents and advice regarding retrovirus generation; and Jim Artenk for performing the EM.

Address correspondence to: Richard Stanton, Section of Medical Microbiology, Department of Infection, Immunity and Biochemistry, School of Medicine, Cardiff University, Tenvus Building, Heath Park, Cardiff CF14 4XN, United Kingdom. Phone: 44.0.29.20687319; Fax: 44.0.29.20742161; E-mail: StantonRJ@cf.ac.uk.

8. Quinlan GV Jr, et al. Comparative virulence and packaging cells (86) using Effectene (QIAGEN) according to manufacturer’s instructions and harvesting the supernatant 48 hours PT. Retrovirus was bound to culture dishes with Retronecim (Clontech) and used to infect HTERT-immortalized HFFs. The cells were selected in puromycin (1 μg/ml) at 48 hours PI.

Inserting tet operators into BACs. Tet operators were inserted by recombineering. To insert 1 tet operator 19 bp upstream of the RL13 coding region, the lacZ/amp/sacB cassette was amplified with primers SacBF-RL13-1 and SacBR-RL13-2 and inserted into the BAC and then removed with oligonucleotide tetO-RL13, leaving behind 1 tet operator. To insert 2 tet operators 33 bp upstream of the UL131A coding region, the lacZ/amp/sacB cassette was amplified with primers SacBF-UL131A and SacBR-UL131A and inserted into the BAC and then removed with oligonucleotide tetO-UL131A, leaving behind 2 tet operators.

Purification of HCMV virions. HCMV particles in cell supernatants were separated into virion, dense body, and noninfectious enveloped particle (NIEP) fractions by positive density/negative viscosity gradient centrifugation as described previously (88). Particle concentrations in the preparations were estimated by counting negatively stained samples by EM in relation to a standard concentration of latex beads. To separate virion envelope proteins from capsid and tegument proteins, 10 pumps were mixed 1:1 with envelope stripping buffer (2% Nonidet-P40 in PBS) and incubated for 15 min at 4°C. Particles were pelleted (12,000 g for 5 minutes at 4°C), and the soluble envelope fraction was harvested. The insoluble capsid/tegument material was washed twice with envelope-stripping buffer and once in PBS before being solubilized in SDS-PAGE sample buffer.

SDS-PAGE and Western immunoblotting. Protein samples were separated by SDS-PAGE using 4%-12% Bis-Tris NuPAGE protein gels (Invitrogen) according to the manufacturer’s instructions, then transferred to nitrocellulose by semi-dry transfer. Membranes were blocked overnight in blocking buffer (3% BSA in PBS containing 0.1% Tween 20 [PBST]) and then incubated with primary antibody for 1 hour. Following washes in PBST, secondary antibody was incubated for 1 hour and washed 3 times. Bound antibody was detected using ECL–Western blotting detection system (RPN 2132; Amersham) and exposure to film. Primary antibodies were chicken anti-V5 (9113; Abcam), mouse anti-gB (CMV-023-40154; Capricorn), and mouse anti-pp65 (CMV-018-48151; Capricorn). Secondary antibodies were goat anti–chicken-HRP (6877; Abcam) and goat anti–mouse-HRP (170-6516; Bio-Rad).

Glycosed treatment. Samples were treated with PNGase F (P0705S; New England BioLabs) or ENDO H (P0703S; New England BioLabs) according to the manufacturer’s instructions. Briefly, the samples were denatured in glycoprotein-denaturing buffer at 100°C for 10 minutes and cooled to 0°C for 5 minutes. The samples were then digested overnight at 37°C with PNGase F or ENDO H before being analyzed by Western immunoblotting.

Immunogold electron microscopy. Purified virions were air-dried to the surface of Formvar-coated EM grids. The grids were treated with chicken anti-V5 antibody (9113) for 4 hours at room temperature, washed 3 times with PBS, and incubated with goat anti-chicken antibody (39604; Abcam) conjugated to 6-nm gold particles for 1 hour at room temperature. After further washing in PBS, the grids were negatively stained with phosphotungstic acid and subjected to EM.


