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


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 passed 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 of the viral genome. The HCMV genome encodes a complex and thereby render the virus incapable of infecting epithelial, endothelial, and certain myeloid cell types (12, 44–48).

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 passed 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.

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 passed 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.

Authorship note: Richard J. Stanton and Katarina Baluchova contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: J Clin Invest. 2010;120(9):3191–3208. doi:10.1172/JCI42955.
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 a BAC containing the complete HCMV genome complement. HCMV strain Merlin (ATCC VR-1590) was selected because it is designated the reference HCMV genome (27). Its expression had not, to our knowledge, previously been characterized. Use of the BACs showed that RL13 encodes a noncoding region 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 U5 and US 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 U5 and U5 present in the standard arrangement was selected for further analysis.

Recombinering (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 recombinering, 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 NheI 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>
<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>CAAAAAACC → CAAAAAAC</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>GGCQAAC → GGCQAAC</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>GACGAGT → GACGAGT</td>
<td>UL36</td>
<td>R – C</td>
</tr>
<tr>
<td>49791</td>
<td>AGGAGTT → AGGAGTT</td>
<td>UL36</td>
<td>P – S</td>
</tr>
<tr>
<td>49792</td>
<td>GGGTTGC → GGGTTGC</td>
<td>UL36</td>
<td>H – Q</td>
</tr>
</tbody>
</table>

194592–194643 (1025–1076) ACACCCCCGTCCCCGACACCCCCGTCCCCGACACCCCCGTCCCCG → G

None (b/b) None

195063 (605) GGAGCGC None (b/b) None

194781  CGGGAGA → CGGGAGA None (b’ only) None

Nt positions are given with respect to NC_006273. f, frameshift. Altered bases underlined.

**Results**

**Cloning the HCMV strain Merlin genome.** We aimed to clone the complete HCMV strain Merlin genome by inserting a BAC vector into a noncoding region 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 U5 and US 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 U5 and U5 present in the standard arrangement was selected for further analysis.

Recombinering (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 recombinering, 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 NheI sites (40 bp in total) located between US28 and US29 (Figure 1).
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 changes 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/c 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/c 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 question of at which stage these mutations had arisen. The original clinical sample, using the same primers and conditions as individual M13 clones were mutated at various different positions from Merlin (p3) was also consistent with this virus stock comprising a mixture of RL13 mutants (Table 2). Thus, 18/22 PCR clones exhibited substitutions that were also detected in the M13 clones. Each of the remaining 4/22 PCR clones contained unique mutations in RL13 (data not shown), though it is not possible to exclude the possibility that these resulted from PCR errors.

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 Merlin sequence was determined from bacteriophage M13 clones derived from Merlin at the pAL1031 stage were mutated in RL13 raised the question of which stage these mutations had arisen. The original Merlin sequence was determined from bacteriophage M13 clones derived from Merlin (p3). Retrospective analysis of the sequence database revealed that, although the RL13 consensus was WT, all individual M13 clones were mutated at various different positions (Table 2). Moreover, an analysis of individual PCR clones produced from Merlin (p3) was also consistent with this virus stock comprising a mixture of RL13 mutants (Table 2). Thus, 18/22 PCR clones exhibited substitutions that were also detected in the M13 clones. Each of the remaining 4/22 PCR clones contained unique mutations in RL13 (data not shown), though it is not possible to exclude the possibility that these resulted from PCR errors.

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 predominantly WT in the clinical sample and are consistent with RL13 mutants having arisen and been selected during cell culture.

**Table 2**

<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>CATGAC → CATCAC</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>ATGCAAT → ATGCAAT</td>
<td>125</td>
<td>W → C</td>
<td>2/8</td>
<td>1/22</td>
</tr>
<tr>
<td>1163</td>
<td>ATGCAAT → ATGCAAT</td>
<td>125</td>
<td>W → C</td>
<td>2/8</td>
<td>4/22</td>
</tr>
<tr>
<td>11900</td>
<td>ATTGAAA → ATTCAAA</td>
<td>238</td>
<td>E → Q</td>
<td>2/8</td>
<td>4/22</td>
</tr>
<tr>
<td>11884</td>
<td>ATATGGA → ATATTGG</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.)

None of the BAC or PCR clones yielded a sequence representing the WT RL13 coding region.

The finding that all genomes in Merlin (p5) and BACs derived from Merlin at the pAL1031 stage were mutated in RL13 raised the question of which stage these mutations had arisen. The original Merlin sequence was determined from bacteriophage M13 clones derived from Merlin (p3). Retrospective analysis of the sequence database revealed that, although the RL13 consensus was WT, all individual M13 clones were mutated at various different positions (Table 2). Moreover, an analysis of individual PCR clones produced from Merlin (p3) was also consistent with this virus stock comprising a mixture of RL13 mutants (Table 2). Thus, 18/22 PCR clones exhibited substitutions that were also detected in the M13 clones. Each of the remaining 4/22 PCR clones contained unique mutations in RL13 (data not shown), though it is not possible to exclude the possibility that these resulted from PCR errors.

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 predominantly WT in the clinical sample and are consistent with RL13 mutants having arisen and been selected during cell culture.

**Growth properties of virus generated from the Merlin BAC.** Virus (RCMV1111) was readily recovered from HFFFs transfected with pAL1111 (Figure 1) and exhibited a restriction endonuclease profile identical to that of Merlin (p5) (Figure 3A). A pattern diagnostic of the presence of all 4 genome isoforms was distinguishable in the linear RCMV1111 genome. Infections of HFFFs at low MOI (0.01 PFU/cell) demonstrated that RCMV1111 grew with kinetics comparable to those of Merlin (p5) (Figure 3B). In generating RCMV1111, the BAC vector was excised from the genome to leave only a residual sequence of 40 bp between 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 3C). 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 seamless-ly by recombineering, both singly and in combination, culminating in a set of 4 BACs (pAL1111, pAL1119, pAL1120, and pAL1128; Table 3). The complete sequence of the fully repaired BAC (pAL1128) was verified by Sanger 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 further, HFFFs transfected with the eGFP-tagged BACs without overlay were monitored over time. Plaques formed initially by viruses with UL128, RL13, or both genes repaired (RCMV1159, RCMV1160, and RCMV1161) were subsequently outgrown by uninfected cells. To encourage virus dissemination, cells transfected with pAL1159, pAL1160, or pAL1161 were trypsinized at 7-day intervals and reseeded in the same flask without the addition or removal of cells. The eGFP reporter gene was used to identify infected cells (Figure 5B), and cell-free virus levels were measured by titration (Figure 5C). The rates of virus spread through the monolayer (Figure 5B) were consistent with the

(groups 1, 2, and 4) in BACs, plus 3 additional groups (classes 5–7).
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 UL128L.

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 fibroblasts.
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 or pAL1161 (Table 3) and the size of plaques was measured (Figure 6A). At all time points measured, plaques from the virus lacking RL13 (pAL1160) were larger than those from the virus in which RL13 had been repaired (pAL1161). Therefore, RL13 conferred a growth restriction in epithelial cells as well as fibroblasts. In contrast, UL128L is known to confer a restriction only in fibroblasts.

Cell cultures transfected with the eGFP-tagged BACs were monitored simultaneously for the number of infected cells (Figure 6B) and amount of infectious virus released into the supernatant (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 HFFs, 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 HFFs, and, unlike HFFs, 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 RCVM1161 (Figure 7B). The 55-kDa proteins were susceptible to EndoH digestion, indicative of them being an ER-retained immature form, whereas the 80- and 100-kDa proteins were resistant to EndoH digestion.

Rapid generation and selection of RL13 and UL128L mutants in fibroblasts. To generate working stocks of BAC-derived viruses in HFFs, 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 HFF 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 passed 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 adenovirus (RAd) vector. The V5 epitope was also fused to RL13 within the strain Merlin BACs, in both a UL128+ (RCMV1279) and a UL128− (RCMV1280) background (Table 3). The primary translation product of RL13 is predicted to be a 35-kDa protein containing a signal sequence, a transmembrane domain, 7 potential N-linked glycosylation sites, and 26 potential O-linked glycosylation sites. When RL13V5 was expressed in isolation using the RAd vector, 80-kDa and 55-kDa species were detected (Figure 7A), whereas 100-kDa and 55-kDa species were detected in cells infected with RCMV1279 (RL13V5:UL128−; Figure 7B). The 55-kDa proteins were susceptible to EndoH digestion, indicative of them being an ER-retained immature form, whereas the 80- and 100-kDa proteins were resis-

### 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>
<tr>
<td>BACs tagged with eGFP</td>
<td></td>
<td></td>
<td></td>
</tr>
<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>
<tr>
<td>BACs containing RL13 tagged with a V5 epitope</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pAL1279</td>
<td>RCMV1279</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pAL1280</td>
<td>RCMV1280</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BACs tagged with eGFP, tet Operators before RL13/UL128L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pAL1393</td>
<td>RCMV1393</td>
<td>–</td>
<td>+ (tetO)</td>
</tr>
<tr>
<td>pAL1448</td>
<td>RCMV1448</td>
<td>+ (tetO)</td>
<td>–</td>
</tr>
<tr>
<td>pAL1498</td>
<td>RCMV1498</td>
<td>+ (tetO)</td>
<td>–</td>
</tr>
</tbody>
</table>

**Figure 4**

Growth of BAC-derived viruses in HFFs. Images show eGFP expressed by plaques formed in HFFs 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 the non-tet–controlled virus (RCMV1393) (Figure 9, B and C). Similarly, 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 HFFF-tet cells with a tet-controlled UL131A virus, RCMV1393 (Figure 9, D and E). Titer 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 UL128 from RCMV1393 and RCMV1448, respectively, showed all contained the WT sequence. Following passage in HFFs with this RCMV1393 stock, virus was able to infect ARPE19 cells and produced plaques of a size equivalent to the non–tet-controlled parental virus RCMV160 (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 RL13 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 (identified in the present study). To date, HCMV research has by necessity used virus-like particles (VLPs) 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 clin-
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></td>
<td>Mutation</td>
<td>Position</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>^A</td>
<td>11363</td>
</tr>
<tr>
<td>1</td>
<td>G → T</td>
<td>11901</td>
</tr>
<tr>
<td>2</td>
<td>ΔA</td>
<td>11363</td>
</tr>
<tr>
<td>3</td>
<td>^A</td>
<td>11363</td>
</tr>
<tr>
<td>4</td>
<td>A → G</td>
<td>11297</td>
</tr>
<tr>
<td>5</td>
<td>C → T</td>
<td>11427</td>
</tr>
<tr>
<td>6</td>
<td>A → G</td>
<td>11956</td>
</tr>
<tr>
<td>RL13-UL128- BAC transfected into HFFFs</td>
<td>C → T</td>
<td>176229</td>
</tr>
<tr>
<td>7</td>
<td>T → C</td>
<td>176845</td>
</tr>
<tr>
<td>8</td>
<td>CAAGA → TCTTG</td>
<td>176915-9</td>
</tr>
<tr>
<td>9</td>
<td>ΔT</td>
<td>176456</td>
</tr>
<tr>
<td>RL13-UL128- transfected into HFFFs</td>
<td>None</td>
<td>176917</td>
</tr>
<tr>
<td>10</td>
<td>None</td>
<td>11363</td>
</tr>
<tr>
<td>11</td>
<td>^A</td>
<td>11363</td>
</tr>
<tr>
<td>12</td>
<td>ΔT</td>
<td>176456</td>
</tr>
<tr>
<td>13</td>
<td>ΔT</td>
<td>176456</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.

critical sample that gave rise to Merlin. Mutations that affected HCMV coding regions were sequentially repaired by recombinase, so that the only relic in virus recovered from BACs by DNA transfection was a 40-bp sequence containing a loxp and an Nhel site located between the US28 and US29 coding regions. The presence of this sequence did not affect expression of the flanking genes.

The high level of sequence identity between pAL1053 and strain Merlin (p3) demonstrated that the HCMV genome had suffered minimal perturbation during BAC cloning. In pAL1053, non synonymous mutations in protein-coding regions were identified in UL36, UL128, and RL13. The differences in UL36 were specific to pAL1053 and its predecessor, pAL1031, and were presumably derived from a small proportion of mutants that arose in cell culture. HCMV and rhesus cytomegalovirus strains have previously been reported in fibroblasts after p20 (44), p22 (45), or p15–p20 (26). When the lesion was repaired in the BACs, alternative defects were selected when clinical isolates are passaged in fibroblasts (12, 26, 27, 44, 45, 47, 48). Merlin was originally selected for detailed investigation from a series of clinical isolates partly on the basis of its congenitally infected neonates (as was the case for the sample from which Merlin was isolated), and this indicates that conditions exist in vivo that are compatible with efficient replication of WT HCMV. It is possible that RL13 regulates a switch to productive HCMV replication in vivo.

The UL128 locus is essential for the efficient infection of epithelial 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 investigation from a series of clinical isolates partly on the basis of its excellent growth properties in fibroblasts. However, this process favored the selection of a UL128 mutant, which emerged during p1 (26). When the lesion was repaired in the BACs, alternative defects in the UL128 locus emerged by p3–p4. In other studies monitoring endothelial cell tropism in clinical isolates, loss of phenotype has been reported in fibroblasts after p20 (44), p22 (45), or p15–p20 (28). Such studies defined a virus “passage” as corresponding to...
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 HFFs infected with the RAd expressing RL13V5 (×640 magnification). (D and E) Immunofluorescence performed for the indicated antigens either on HFFs 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 10^9 particles were used in the initial infection of a 25-cm^2 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 10^9 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 HFFFs, 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 mutated viruses; however, this may simply reflect the slower growth of the WT virus.

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 among all human herpesviruses, 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, 39, 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 HFFs 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 HFFs (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 that have lost RL13 function. When the low passage HCMV strain Toledo was evaluated in human virus challenge, mild to severe disease was induced in immunocompetent individuals (both seronegative and seropositive) receiving doses ranging between 10 and 1,000 PFU (9). Indeed, the capacity of this strain to induce clinical disease was impressive. However, Toledo is known to have a large, in-frame deletion in RL13 and an ablation of UL128 (27). The absence of RL13 function as a repressor of HCMV replication now DNA polymerase, and religated to remove the existing cos site underlined, CATACATTATACGAAGTTAT

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

Cells and viruses. HFFFs 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 by plaque assay for 14 days on HFFs using a 1% Avicel overlay (79). Cultures that exhibited small plaques at 14 days PT were incubated for a further 14 days and recounted. Plaques were visualized using an ORCA-ER camera mounted on a Leica DMIRBE microscope, and sizes were computed using Openlab 3 software (Improvision).

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 (ATCCTGAACATGAAGACTCGTGT) and RL13R (GAATAAACA-

Figure 10
Simultaneous repression of UL128L and UL13 by BAC-derived viruses in HFFF-tet cells. Plaque sizes of viruses generated at 3 weeks PT in HFFFs (A) or HFFF-tet (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 Pael and inserted into Pacl-digested pAL185, 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 HFFs. 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 pAL1033 (Figure 1). To generate the insert, 2 PCR products were produced, digested with Nhel, and ligated together. A 482-bp region encompassing the loxP site adjacent to US34 in pAL1026 was amplified using primers LoxP-R (GGCCGCGTACATAATGTGATATGCTTTGATGC-TATAGGAAGTTATGCATGCAAGCTTGAGTATTC; NheI site underlined) and Chlor-F (GGCCAGATCTGTCGCGAAGATGCTTAATGGA; BglII site underlined). A 5.2-kb region of the Merlin genome spanning genes US34 to US34A was amplified using primers US29F (GGCCGCTAGGACCCTGGCGCTTTTCATACAA; Nhel site underlined) and US34AR (GGCCGAATTCATCAATCTGGGACTTTCATCCTGCATTA; BglII site underlined). Each amplicon was cloned into pCR2.1-TOPO (Invitrogen) for sequencing. Two different amplicons in each amplion 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 with US29-US34A and generating pAL1040 (Figure 1). The eGFP/Puro cassette in pAL1040 was replaced by a Cre recombinase 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-GFPFPuroHom and SacBR-GFPFPuroHom and inserted into pAL1040, thus generating pAL1047 (Figure 1). The Cre recombinase expression cassette was amplified from YD-C66 (a gift from D. Yu; ref. 53), using primers CreF-US34OHom 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-US36Hom and SacBR-US36Hom and inserted in place of UL36. UL36 was then amplified from Merlin DNA using primers UL36F and UL36R and ligated together 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-JE2Hom and SacBR-JE2Hom. In the resulting BACs, the lacZ/amp'/sacB cassette was replaced by an IRES-eGFP cassette amplified from pIRES2-GFP (Clontech) using primers IRESF-JE2Hom and GFPRI-JE2Hom.

**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 pAL1128 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 (GATCAGTGGAGCAGCTTTCA and US29R (CTTGGGCTTCCAGACCGG) and UL123 (IE2) transcripts were amplified using primers IE2F (CGGAGAAGATGGAACAGATGATGA and IE2R (TCTCCTAGCACCCTCCTC). 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 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. Hopken; ref. 84) was used (1:5) in combination with goat anti-mouse Alexa Fluor 488 (1:200) for detection (Invitrogen). HCMV infection was monitored directly by eGFP expression or indirectly by downregulation of MHC class I from the cell surface using antibody W6/32 (1:100) followed by 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 HFFF-tet cell line.** The coding region of the tet represor was amplified using primers tetR-F (GGGCCGGATTCCCTTAGGCC-GAATGATATGCTCA; BamHI site underlined, initiation codon in bold) and tetR-R (GGCCGAATTCATCAATCTGGGACTTTCATCCTGCATTA; BglII site underlined). The amplified fragment was cloned into the BamHI/EcoRI sites of the retrovirus vector pMXs-IP (85). Retrovirus stocks were produced by transfecting 293Phoenix
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 Retronectin (Clontech) and used to infect HEK293-IMMART-1cells. 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-RL13-2 and SacBR-RL13 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.

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 after the RL13 ORF within the BAC; then the cassette was removed using overlapping oligos V5-RL13-F and V5-RL13-R, which fused the coding sequence for a V5 tag to the tail of gene RL13.

Generating replication-deficient RAd vectors. RL13 was cloned by recombineering into a RAd using the AdZ 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 preparations were determined by absorbance at 260 nm.

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 incubated for 15 minutes at 4°C. Samples were probed with primary antibody for 1 hour and washed again, and mounted. Primary antibodies were goat anti-RV (9113, Abcam), sheep anti-human TGN46 (AHP500; Sero-tec), mouse MoAb anti-pp65 (6502, Abcam), and mouse anti-gB (2470-5437, Biogenes). Secondary antibodies were donkey anti-rabbit Cy3 (AP194C; Millipore), donkey anti-antibody IgG–Alexa Fluor 633 (A21100; Invitrogen), and donkey anti-mouse IgG–Alexa Fluor 488 (A-2102; Invitrogen). The intracellular localizations of antibody-tagged proteins were examined under laser illumination in a Zeiss LSM 510 confocal microscope, and images were captured using LSM software.

Statistical analysis was performed using GraphPad Prism 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; Neil Copeland, Dong Yu, and Uta Höpken for reagents; Richard Darley and Philip Taylor for reagents and advice regarding retrovirus generation; and Jim Atken for performing the EM.

Received for publication March 12, 2010, and accepted in revised form June 23, 2010.

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


8. Quinnan GV Jr, et al. Comparative virulence and
immunogenicity of the Towne strain and a nonat- 
1984;100(8):478–483.
9. Plotkin SA, Starr SE, Friedman HM, Gonzol E, 
Weibel RE. Protective effects of Towne cytome-
glovirus vaccine against low-passage cytome-
10. Elek SD, Stern H. Development of a vaccine against 
mental retardation caused by cytomegalovirus 
11. Chu TA, Taylor K, Noble GW, Duke GM, Mocar-
ski ES, Spaete RR. Human cytomegalovirus clinical 
isolates carry at least 19 genes not found in labora-

12. Prichard MN, Penfold ME, Duke GM, Spaete RR, 
Kemble GV, et al. Genetic content of wild-type 
genome revisited: comparison with the chimpan-
14. Skalaktsaya A, Bartle LM, Chistendem T, 
McCor-
mick AL, Mocarski ES, Goldmacher V. A cytome-
glovirus-encoded inhibitor of apoptosis that sup-
presses caspase-8 activation. Proc Natl Acad Sci U S A.
15. Dargan DJ, Jamieson FE, MacLean J, Dolan A, Addi-
son CM, Geocheok DJ. The DNA published sequence 

data of human cytomegalovirus strain AD169 lacks 929 
Reassessing the organization of the UL42-UL43 
region of the human cytomegalovirus strain AD169 
17. Brown JM, Kaneshima H, Mocarski ES. Dramatic 
interstrain differences in the replication of human 
analysis of variants of human cytomegalovirus 
89(pt 10):2375–2380.
19. Penfold ME, H. Cytomegalovirus encodes a 

depot alpha chemokine. Proc Natl Acad Sci U S A.
receptor superfamily member in virulent strains 
a human class I-like molecule (UL142) that func-
175(11):7457–7465.
cell-activating ligand CD155 by human cytome-
23. Ceroni C, et al. Human cytomegalovirus strain-
dependent changes in NK cell recognition of infect-
T. Human cytomegalovirus sequences expressed in 

elated-infected individuals promote a latent infec-
25. Petrucelli A, Rak M, Grainger L, Goodrum F. Char-
acterization of a novel Golgi apparatus-localized 
lateness determinant encoded by human cytome-
26. Akter P, et al. Two novel spliced genes in human cyto-
27. Dolan A, et al. Genetic content of wild-type 
with adaptation of human cytomegalovirus to growth 
29. Chandler SH, Handsfield HH, McDougall JK. Isol-
ation of multiple strains of cytomegalovirus from 
Women attending a clinic for sexually transmitted 
30. Collier AC, Chandler SH, Handsfield HH, Corey L, 
McDougall JK. Identification of multiple strains of 
31. Spector SA, Hirata KK, Newman TR. Identification 
of multiple cytomegalovirus strains in homosexual 
men with acquired immunodeficiency syndrome. 
32. Baldanti F, et al. Coinfection of the immunocom-
promised but not the immunocompetent host by 
multiple human cytomegalovirus strains. Arch 
multiple cytomegalovirus strains from individual 
34. Drew WL, Sweet ES, Miner RC, Mocarski ES. Mul-
tiple infections by cytomegalovirus in patients with 
35. Adler B, Scrivano L, Ruzcics Z, Rupp B, Sinzger C, 
Koszinowski U. Quantification of replication of clinical 
cytomegalovirus gH/gL/UL128-131 complex that 
mediates entry into epithelial and endothelial cells. 
36. Bykman BJ, Jarvis MA, Drummond DD, Nelson 
JA, Johnson AD. Human cytomegalovirus infects 
endothelial and epithelial cells depends on 
genes UL128 to UL150 and occurs by endocytosis 
cytomegalovirus gH/gL/UL128-131 complex that 
mediates entry into epithelial and endothelial cells. 
38. Wang D, Shenk T. Human cytomegalovirus UL131 
open reading frame is required for epithelial cell 


target fusion derivatives, regulates inducible gene 
expression in mammalian cells. Hum Gene Ther.
40. Ryckman BJ, Chase MC, Johnson DC. HCMV gH/

gL interaction during entry into endothelial cells. 
endotheliotropic virus strain derived from human 
cytomegalovirus TB40/E. J Gen Virol. 2008;89(pt 2):
339–356.
42. Copeland NG, Jenkins NA, Court DL. Recom-
bining: a powerful new tool for mouse functional 

43. Court DL, Sawczuk IS, Court JC. Genetic 
engineering using homologous recombination. 
44. Smith GA, Enquist LW. A self-recombinating 
cytomegalovirus that directs the viral latency 

74(7):7720–7729.
45. Cunningham C, et al. Sequences of complete human 
cytomegalovirus genomes from infected cell 
cultures and clinical specimens. J Gen Virol. 2007;
46. Sanchez V, Clark CL, Yen JY, Dwarkanath R, 
Spector DH. Viable human cytomegalovirus recombi-
nants with an internal deletion of the gB gene 

47. Griffin C, et al. Characterization of a highly gly-
cosylated form of the human cytomegalovirus 
HCA class I. Ecolab, Inc. coli: a new approach for 
analysis of herpesvirus pathogenesis. Proc Natl Acad Sci 
48. Hobom U, Brunow M, Messerer M, Hahn G, 
Koszinowski U. Fast screening procedures for random transposition vectors in human herpesvirus genomes: mutational analysis of human cytome-
49. Cunningham C, et al. Sequences of complete human 
cytomegalovirus genomes from infected cell 
cultures and clinical specimens. J Gen Virol. 2007;