SM6: ORF-disrupting mutations: additional discussion

Low-frequency mutations. We cannot rule out the possibility that other genes are mutated in our isolates. If mutations are present at different sites in different individual clones in the population, the consensus sequence could still represent the intact ORF sequence, while, in fact, the majority of clones contain an ORF-disrupting mutation, as is the case for the RL13 gene of strain Merlin (1). On the other hand, a minority of clones could still harbor an intact ORF sequence when the consensus sequence shows a mutation as in RL13 of strain BE/8/2012. Because of the lack of connectivity of short-read NGS data, however, it cannot be excluded that a wild-type variant at the consensus mutation site is accompanied by a mutation at another site. Thus, although we can never exclude the presence of wild-type gene variants at low frequencies, at least for the mutations that were confirmed in the original clinical isolate by direct sequencing of PCR products, we show that mutated gene variants also predominate in the clinical isolate, indicating their viability in the host.

Mutations in RL13 and UL128. Experiments studying the effects of cell culture passaging on HCMV genes have identified RL13 and one of the UL128 locus genes UL128, UL130 and UL131A as the first genes that are mutated because of their inhibitory effects on HCMV replication in fibroblast cells (1, 2). When evaluating these genes in our 96 isolates, we identified only two obvious disruptive mutations in RL13 (BE/7/2012 and BE/8/2012) and UL128 (BE/11/2011 and BE/2/2013), respectively (Supplemental material 2). Unfortunately, we were unable to confirm these mutations in 3 out of 4 original isolates due to insufficient sample volumes. For strain BE/11/2011, however, a C to T substitution was noticed in the first splice-donor site of gene UL128 after passaging, which might affect proper splicing of the first intron. Upon inspection of the Sanger sequencing traces, we did not identify the presence of this substitution in the original clinical isolate. Interestingly, strain 3157 contains a distinct substitution (C to G) at the same site, which the authors also suggested to be an artifact of culture passaging (3). In the RL13 gene of strain BE/8/2012, 45% of reads still contain the two nucleotides that are deleted and cause premature termination in the consensus. Reciprocally, for strain BE/2/2013, 48% of reads contain a substitution in the start codon, which completely omits the RL13 gene.

Such 50/50 cases of wild-type and mutant variants were not observed for other genes and are probably a sign of active selection that favors RL13 mutants during replication in fibroblasts. Together, it appears that mutations in genes RL13 and UL128 are highly probable to be artifacts of culture passaging. Apart from the confirmed passage artifact in UL128 of BE/11/2011, we did keep these mutations listed in Table 5 and Additional file 6, as we can never be 100% certain that these were not present in the original isolates without sequencing the clinical material directly.

Further discussion of genes containing disruptive mutations. Mutations in genes UL142 and US27 are rare, but were confirmed in the original clinical material. Similar to the UL40 signal peptide, pUL142 counteracts NK cell recognition of infected cells missing normal MHC presentation, but reciprocally to the UL40 peptide, it does so by downregulating the surface expression of NK cell activating ligand MICA (4). US27 encodes a G protein-coupled receptor homolog associated with a role in extracellular spread (5), enhancement of signaling through host chemokine receptor CXCR4 (6) and promotion of cell proliferation and survival (7). For genes UL30 and UL150, mutations were also confirmed, but functional information is currently unavailable. Mutations in five other genes are rare, could not be analyzed in the original isolates and might be artifacts of cell culture passaging. These include genes UL133 and UL136 that regulate replication in a cell type - dependent manner (8, 9) and IRS1 that functions in inhibiting the protein kinase R antiviral pathway together with TRS1 (10). Genes UL148 and UL150A are poorly characterized.

1. **Stanton RJ, Baluchova K, Dargan DJ, Cunningham C, Sheehy O, Seirafian S, McSharry BP, Neale ML, Davies JA, Tomasec P, Davison AJ, Wilkinson GW.** 2010. Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. J Clin Invest **120:**3191-3208.
2. **Dargan DJ, Douglas E, Cunningham C, Jamieson F, Stanton RJ, Baluchova K, McSharry BP, Tomasec P, Emery VC, Percivalle E, Sarasini A, Gerna G, Wilkinson GW, Davison AJ.** 2010. Sequential

mutations associated with adaptation of human cytomegalovirus to growth in cell culture. J Gen Virol **91:**1535-1546.

3. **Cunningham C, Gatherer D, Hilfrich B, Baluchova K, Dargan DJ, Thomson M, Griffiths PD, Wilkinson GW, Schulz TF, Davison AJ.** 2010. Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens. J Gen Virol **91:**605-615.
4. **Ashiru O, Bennett NJ, Boyle LH, Thomas M, Trowsdale J, Wills MR.** 2009. NKG2D ligand MICA is retained in the cis-Golgi apparatus by human cytomegalovirus protein UL142. J Virol **83:**12345-12354.
5. **O'Connor CM, Shenk T.** 2011. Human cytomegalovirus pUS27 G protein-coupled receptor homologue is required for efficient spread by the extracellular route but not for direct cell-to-cell spread. J Virol **85:**3700-3707.
6. **Arnolds KL, Lares AP, Spencer JV.** 2013. The US27 gene product of human cytomegalovirus enhances signaling of host chemokine receptor CXCR4. Virology **439:**122-131.
7. **Tu CC, Spencer JV.** 2014. The DRY Box and C-Terminal Domain of the Human Cytomegalovirus US27 Gene Product Play a Role in Promoting Cell Growth and Survival. PLoS One **9:**e113427.
8. **Umashankar M, Petrucelli A, Cicchini L, Caposio P, Kreklywich CN, Rak M, Bughio F, Goldman DC, Hamlin KL, Nelson JA, Fleming WH, Streblow DN, Goodrum F.** 2011. A novel human cytomegalovirus locus modulates cell type-specific outcomes of infection. PLoS Pathog **7:**e1002444.
9. **Caviness K, Cicchini L, Rak M, Umashankar M, Goodrum F.** 2014. Complex Expression of the UL136 Gene of Human Cytomegalovirus Results in Multiple Protein Isoforms with Unique Roles in Replication. J Virol **88:**14412-14425.
10. **Hakki M, Marshall EE, De Niro KL, Geballe AP.** 2006. Binding and nuclear relocalization of protein kinase R by human cytomegalovirus TRS1. J Virol **80:**11817-11826.