

Reply to Jensen and Kowalik: Consideration of mixed infections is central to understanding HCMV intrahost diversity.

Charlotte Houldcroft^{*1}, Juli Cudini^{*2}, Richard Goldstein³, Judith Breuer^{**3}

*These authors contributed equally

** Corresponding author Division of Infection and Immunity, University College London. The Cruciform Building, Gower street, WC1E 6BT, UK.

Tel: 44-203 108 2030

Email: j.breuer@ucl.ac.uk

1. Department of Medicine, Cambridge University, Addenbrookes Hospital, Box 157, Hills Rd, Cambridge CB2 0QQ, UK.
2. Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK.
3. Division of Infection and Immunity, University College London, The Cruciform Building, Gower Street, WC1E 6BT, UK.

Kowalik and Jensen (1) have reported that intra-host variation in HCMV approaches levels similar to those of HCV, with fast mutation rates mooted as one explanation (2). While we discussed that HCMV mutation rates were postulated as an explanation for high diversity, the focus of our work is on observed inconsistencies in nucleotide diversity between and within patients (3). Kowalik and Jensen did calculate HCMV mutation rates to be similar to MCMV but maintained that this could underestimate the true levels (2). In contrast, our study showed that in the absence of mixed infections, HCMV is no more diverse than other DNA viruses, and considerably less so than chronic RNA viruses. This simple conclusion is different to that of Kowalik and Jensen and had not been stated prior to our publication.

Their previous work concluded that diversity was similar in mixed and single HCMV infections and correlated with selection (4). Yet in the presence of pervasive recombination or mixed infection, calculation of selection is unreliable and likely to be upwardly biased. Their use of consensus sequence principle component analysis (PCA) to identify mixed and single infections is also not optimal. PCA clusters by polymorphisms, ignoring haplotypes, and thus identifies mixed infections only where the constituent viruses are highly divergent, under-estimating their true number. We did not cite their work as we did not have the space to adequately discuss the discrepancies in our approaches.

We agree that HCMV compartmentalization can exist eg. within the vitreous humor of the eye (5). However, our analyses were conducted on HCMV from blood where Kowalik and Jensen argue diversity is highest, so it is unlikely that our lower diversity estimates were an artefact of compartmentalization (6). Kowalik and Jensen also mention the importance of bottlenecks in generating diversity, arguing that

severe bottlenecks can both decrease (1,5) and increase (2) diversity depending on the environment. Where we observe high diversity, we detect both haplotypes from the earliest timepoint, indicating that the initial infection bottleneck had already occurred. We agree that recombination renders phylogenetic (and indeed selection) analyses inaccurate. The RL11 region has the advantage of being in stronger linkage disequilibrium (7) and thus less affected by recombination. Notwithstanding, Figure 4 in our paper clearly demonstrates that the pairwise differences we detected between haplotypes is not limited to the RL11D region but extends to multiple regions across the genome (3).

To summarize, Kowalik and Jensen postulate multiple complex models, which have themselves evolved over the years, to explain HCMV intrahost diversity (4, 6, 8). In contrast, we observe the measure of HCMV intrahost diversity can be explained by mixed infections involving genetically distant viral strains. We do not exclude the contribution of mutation or recombination to HCMV evolution over short or longer periods in our PNAS paper (3) or in our previous publications (7, 9). Importantly, our methods can be applied to other pathogens and our conclusions for HCMV are now supported by the findings of other leaders in the field of HCMV genetics (5, 10).

References

1. Kowalik TF, Jensen JD (2019) A consideration of within-host human cytomegalovirus (HCMV) genetic variation. *Proc Natl Acad Sci*.
2. Renzette N, Bhattacharjee B, Jensen JD, Gibson L, Kowalik TF (2011) Extensive Genome-Wide Variability of Human Cytomegalovirus in Congenitally Infected Infants. *PLoS Pathog* 7(5). doi:10.1371/journal.ppat.1001344.
3. Cudini J, et al. (2019) Human cytomegalovirus haplotype reconstruction reveals high diversity due to superinfection and evidence of within-host recombination. *Proc Natl Acad Sci*:201818130.
4. Renzette N, et al. (2015) Limits and patterns of cytomegalovirus genomic diversity in humans. *Proc Natl Acad Sci U S A* 112(30):E4120-8.
5. Hage E, et al. (2017) Characterization of Human Cytomegalovirus Genome Diversity in Immunocompromised Hosts by Whole-Genome Sequencing Directly From Clinical Specimens. *J Infect Dis* 215(11):1673–1683.
6. Renzette N, et al. (2013) Rapid Intrahost Evolution of Human Cytomegalovirus Is Shaped by Demography and Positive Selection. *PLoS Genet* 9(9). doi:10.1371/journal.pgen.1003735.
7. Lassalle F, et al. (2016) Islands of linkage in an ocean of pervasive recombination reveals two-speed evolution of human cytomegalovirus genomes. *Virus Evol* 2(1).
8. Sackman AM, Pfeifer SP, Kowalik TF, Jensen JD (2018) On the Demographic and Selective Forces Shaping Patterns of Human Cytomegalovirus Variation within Hosts. *Pathog (Basel, Switzerland)* 7(1). doi:10.3390/pathogens7010016.
9. Houldcroft CJ, et al. (2016) Detection of Low Frequency Multi-Drug Resistance and Novel Putative Maribavir Resistance in Immunocompromised Pediatric Patients with Cytomegalovirus. *Front Microbiol* 7(SEP):1317.
10. Suárez NM, et al. (2019) Multiple-Strain Infections of Human Cytomegalovirus with High Genomic Diversity are Common In Breast Milk from HIV-Positive Women in Zambia. *J Infect Dis*. doi:10.1093/infdis/jjz209.