Human cytomegalovirus extensively re-organizes the human genome, diminishing TEAD1 transcription factor activity

Khund Sayeed[1,#], Sreeja Parameswaran[1,#], Matthew J. Beucler[2], Lee E. Edsall[1], Andrew VonHandorf[1], Audrey Crowther[1,3], Omer Donmez[1], Matthew Hass[1], Scott Richards[1], Carmy Forney[1], Jay Wright[2], Merrin Man Long Leong[4], Laura A. Murray-Nerger[4,5,6], Ben E. Gewurz[4], Kenneth M. Kaufman[1,7,8] , John B. Harley[8], Bo Zhao[4], William E. Miller[2], Leah C. Kottyan[1,7,9,**], Matthew T. Weirauch[1,7,9,10,11,*]

1. Center for Autoimmune Genomics and Etiology, Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, 45229, USA.
2. Department of Molecular Genetics, Biochemistry & Microbiology, University of Cincinnati, Cincinnati, OH, 45229, USA.
3. Immunology Graduate Program, University of Cincinnati College of Medicine, Cincinnati, OH, 45229, USA.
4. Department of Medicine, Division of Infectious Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 02115, USA.
5. Department of Microbiology, Harvard Program in Virology, Harvard Medical School, Boston, MA, 02115, USA.
6. Center for Integrated Solutions to Infectious Diseases, Broad Institute of Harvard and MIT, Cambridge, MA, 02142, USA.
7. Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, 45229, USA.
8. Research Service, Cincinnati VA Medical Center, Cincinnati, OH 45229, USA.
9. Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, 45229, USA.
10. Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, 45229, USA.
11. Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, 45229, USA.

#These authors contributed equally.

* Lead corresponding author: Matthew.Weirauch@cchmc.org.

** Corresponding author: Leah.Kottyan@cchmc.org

Summary

Human cytomegalovirus (HCMV) infects up to 80% of the world's population. Here, we show that HCMV infection leads to widespread changes in human chromatin accessibility and chromatin looping, with hundreds of thousands of genomic regions affected 48 hours after infection. Integrative analyses reveal HCMV-induced perturbation of Hippo signaling through drastic reduction of TEAD1 transcription factor activity. We confirm extensive concordant loss of TEAD1 binding, active H3K27ac histone marks, and chromatin looping interactions upon infection. Our data position TEAD1 at the top of a hierarchy involving multiple altered important developmental pathways. HCMV infection reduces TEAD1 activity through four distinct mechanisms: closing of TEAD1-bound chromatin, reduction of YAP1 and phosphorylated YAP1 levels, reduction of TEAD1 transcript and protein levels, and alteration of *TEAD1* exon-6 usage. Altered TEAD1-based mechanisms are highly enriched at genetic risk loci associated with eye and ear development, providing mechanistic insight into HCMV's established roles in these processes.

Keywords: virology; functional genomics; human cytomegalovirus; hippo pathway; chromatin looping; chromatin immunoprecipitation sequencing (ChIP-seq); gene regulation; chromatin accessibility; eye development; hearing loss

Introduction

Human cytomegalovirus (HCMV) is nearly ubiquitous in the human population. While primary infection of immunocompetent individuals is typically mild or asymptomatic, the virus can cause significant disease in developing neonates and immunocompromised individuals (Rafailidis et al., 2008; Swanson and Schleiss, 2013). HCMV establishes a lifelong latent infection following primary infection that can periodically reactivate throughout the patient's lifetime, leading to shedding of infectious virus or promotion of virus-associated morbidity in immunocompromised individuals (Forte et al., 2020). HCMV can infect a broad range of cells including fibroblasts, epithelial cells, and hematopoietic progenitor stem cells (Sinzger et al., 2008). Fibroblasts and epithelial cells are typically supportive of a lytic infection, while hematopoietic cells are the reservoir of latent or persistent infection. Fibroblasts have been widely used to study the cellular changes that occur in response to viral infection and for viral propagation in cell culture (Fortunato, 2021; Sinzger et al., 2008; Sinzger et al., 1995). Like all other viruses, HCMV is an obligate intracellular pathogen that relies on the host cell machinery for its replication, assembly, and egress. These complex host-pathogen interactions vary among different cell types, leading to viral infection manifesting itself in primarily a lytic or latent fashion (Lee and Grey, 2020; Rothenburg and Brennan, 2020). Infection of permissive cells is accompanied by widespread changes in host cell gene expression. Indeed, several previous studies have examined pathways altered by HCMV infection of *in-vitro* cultured cells at the level of gene expression (Hein and Weissman, 2021; Lee and Grey, 2020; Li and Kamil, 2016; Van Damme et al., 2016; van Den Pol et al., 1999).

Regulation of gene expression involves intricate interplay between the binding of transcription factors and cofactors, chromatin accessibility, histone marks, and long-range chromatin looping interactions (Lambert et al., 2018). A hallmark feature of viruses is their dependence on the host nuclear environment and host transcriptional processes. For example, some viruses, such as Epstein-Barr virus and Kaposi sarcoma-associated herpesvirus extensively alter human chromatin architecture in order to replicate (Campbell et al., 2022; Jiang et al., 2017). HCMV uses host-derived histones to chromatinize its genome for temporal regulation of the Immediate Early protein IE1 (Zalckvar et al., 2013). However, the mechanisms affecting HCMV driven regulation of host gene expression are not well defined, thus limiting our understanding of the complex interplay between virus and host.

To determine the human gene regulatory mechanisms impacted by HCMV, we examined the changes to chromatin accessibility, transcription factor occupancy, chromatin looping, and gene

expression resulting from HCMV infection. We show that HCMV infection in human fibroblasts and retinal epithelial cells induces large-scale global changes in the human chromatin landscape. Accessible chromatin regions that closed upon HCMV infection were highly enriched for predicted TEAD sites and depleted of CTCF (CCCTC-Binding Factor) sites in both human fibroblasts and retinal epithelial cell lines. Chromatin-immunoprecipitation followed by sequencing (ChIP-seq) experiments confirm that TEAD1 binding sites are significantly depleted at sites that become inaccessible with HCMV infection, with concomitant reduction of chromatin looping interactions and H3K27ac levels.

TEAD1 is a downstream effector of the Hippo signaling pathway, which regulates cell proliferation and cell fate to control organ growth and regeneration (Currey et al., 2021; Ma et al., 2019). TEAD1 activity is primarily controlled by its co-activator YAP1 (Ma et al., 2019; Totaro et al., 2018). We show that HCMV-induced loss of TEAD1 binding is mediated by four distinct mechanisms: 1. Extensive closing of human chromatin that is normally occupied by TEAD1; 2. Reduction of YAP1 and pYAP1 protein levels; 3. Reduction of TEAD1 transcript and protein levels; 4. Exclusion of TEAD1 exon 6. Consistent with these observations, pathway enrichment analysis of differentially expressed genes upon HCMV infection reveals extensive perturbation of the Hippo/TEAD signaling pathway. GWAS-based enrichment analysis reveals that these TEAD1 binding loss events are specifically enriched for genetic variants associated with ear and eye development. Collectively, our data provide novel insights into the mechanisms employed by HCMV upon infection of human cells, with important implications in HCMV-induced human growth defects.

Results

Human cytomegalovirus infection extensively re-organizes human chromatin upon infection

Several previous studies have reported extensive changes in human gene expression profiles in HCMV infected cells (Hertel and Mocarski, 2004; McKinney et al., 2014; Nightingale et al., 2018; Nogalski et al., 2019; Oberstein and Shenk, 2017). However, the regulatory mechanisms underlying these vast changes remain largely unknown. Gene expression programs defined by *cis*-acting DNA elements such as enhancers and promoters typically occur in open or accessible regions of chromatin. We thus sought to determine if HCMV infection leads to genome-wide alterations in chromatin structure. To this end, we used ATAC-seq (Corces et al.,

2017) to measure HCMV-induced changes to human chromatin accessibility in two widely used HCMV infection models: human fibroblasts (HS68 cells) and retinal epithelial cells (ARPE-19 cells) (**Figure 1A**). ATAC-seq experiments were performed in duplicate in uninfected conditions or 48 hours post infection (hpi) (**Figure 1B**). For each cell line, we obtained >200,000 peaks in both uninfected and infected cells, with Fraction of Reads in Peaks (FRiP) scores greater than 0.5 and Transcription Start Sites (TSS) enrichment scores >30 (average of 35.1) (**Supplemental Dataset 1**), all of which greatly exceed the recommendations set forth by the ENCODE Project consortium (Consortium, 2012; Hitz et al., 2023; Luo et al., 2020; Meenakshi S. Kagda et al., 2023). Likewise, our ATAC-seq peaks align strongly with relevant publicly available datasets (**Supplemental Dataset 1**), and we observed very strong agreement between replicates, with samples tightly clustering first by cell type and then by infection status (**Supplemental Figure S1**). Collectively, these results highlight the high quality and internal consistency of our datasets.

We next sought to systematically identify regions of the human genome with differential chromatin accessibility between infected and uninfected cells. To this end, we used DiffBind (Ross-Innes et al., 2012) (see Methods), which identified thousands of chromatin regions (ATAC-seq peaks) with statistically significant changes in signal (**Supplemental Dataset 2**). As expected, most accessible chromatin is common to both uninfected and HCMV infected cells in both cell types - we designate such regions 'unchanged' (**Figure 2A** and **2B**). In fibroblasts, 38,651 peaks were unique to uninfected cells, and hence were closed following infection with HCMV (**Figure 2A** and **2C**). Likewise, 49,003 regions were newly accessible in HCMV infected fibroblasts. We also observed large HCMV-dependent changes to chromatin accessibility in retinal epithelial cells (**Figure 2B** and **Figure 2D**). Collectively, these data indicate that HCMV infection of human cells has widespread effects on the human chromatin accessibility landscape.

Regions of open chromatin largely reflect regulatory regions such as enhancers, which can interact with promoters at large genomic distances spanning many megabases through chromatin looping interactions (Rowley and Corces, 2018). To measure the impact of HCMV infection on functional chromatin looping across the human genome, we next performed HiChIP with an antibody against H3K27ac in uninfected and infected fibroblasts (see Methods). Analysis of the resulting HiChIP data revealed 143,882 and 97,815 chromatin looping interactions in uninfected and infected cells, respectively (**Supplemental Dataset 3**). QC analyses using HiC-Pro (Servant et al., 2015) indicate that the data are of high quality: the final

set of unique valid interaction pairs was 67% of the total sequenced pairs for the uninfected cells and 52% for the infected cells. The number of *trans* interactions was 21% of the sequenced pairs for the uninfected cells and 11% for the infected cells (**Supplemental Dataset 4**), similar to or better than the results obtained in the original HiChIP study (Mumbach et al., 2016). Peaks were called from the HiChIP data and compared to peaks called from H3K27ac ChIP-seq data (see Methods), revealing between 82% and 92% HiChIP peak overlap with H3K27ac ChIP-seq peaks (**Supplemental Dataset 4**). Collectively, these results indicate that our HiChIP data are of high quality.

We next used the HiChIP data to identify HCMV-dependent differential chromatin looping events (see Methods), identifying 74,797 loops that open and 120,864 loops that close following HCMV infection (**Supplemental Dataset 3**). Of the 23,018 loops present in both infected and uninfected cells, only 10 are differential at a 2-fold cutoff and a false discovery rate (FDR) < 0.01. Because the number of loops unique to the infection state were so striking, we repeated loop calling and differential analysis with FDR values less than 0.05, 0.01, and 0.001 (**Supplemental Dataset 3**). For all three cutoffs, the percentage of loops specific to infected cells remains at least 70%. We also randomly downsampled the number of input pairs used for calling loops to verify that our results were not due to a difference in read depth (**Supplemental Dataset 3**). For the three smaller subsets of data, the number of unique loops changed by only one percent. HCMV-dependent losses in chromatin accessibility significantly correspond with HCMV-dependent losses of chromatin looping (enrichment: 2.1-fold, adjusted p-value: 9.6×10^{-192}), but not HCMV-dependent gains of chromatin looping (enrichment: 0.6-fold, adjusted p-value: 1) (**Supplemental Dataset 5**). The full quantification of each chromatin looping event and comparisons of events between conditions are provided in **Supplemental Dataset 6**.

Next, we examined HCMV-altered regions of chromatin accessibility for enriched transcription factor binding site (TFBS) motifs (see Methods). We use unchanged accessibility regions as a baseline for motif enrichment analysis, in order to identify motifs specific to chromatin regions that are uniquely accessible in either uninfected or infected cells. In human fibroblasts, all three categories of regions (unchanged, closed with infection, open with infection) were highly enriched for AP-1 motifs (**Figure 3A**, purple dots), reflecting the important role played by this TF family in virtually every cell type regardless of infection status (Lee et al., 1987; Wolf et al., 2023). Intriguingly, we found substantial differences in motif enrichment for other TF families in HCMV-altered chromatin regions. In regions closed upon HCMV infection, CCAAT-enhancer-binding protein (CEBP), E-box (bound by bHLH TFs), and TEA domain (TEAD) binding motifs

were uniquely enriched (**Figure 3A**, left: blue, orange, and maroon dots along the Y-axis, respectively). In contrast, newly closed chromatin in HCMV infected cells was depleted of CTCF binding sites (**Figure 3A**, left: dark green dots along the X-axis), and newly opened chromatin was enriched slightly more strongly for CTCF sites (**Figure 3A**, right: dark green dots approaching the Y-axis).

We next examined transcription factor binding site motif enrichment in retinal epithelial cells. Similar to human fibroblasts, we found that TEAD motifs were specifically enriched in peaks that were closed with infection (**Figure 3B**, left). Unlike in fibroblasts, we did not identify strong enrichment of CEBP or E-box motifs, presumably due to cell-type differences. In addition to differences in enrichment levels, we also observed substantial and consistent changes in the overall predicted occupancy of TEAD TFs and CTCF in both cell types (**Figures 3C and 3D**). Taken together, these results indicate that HCMV infection extensively re-organizes the human genome in both fibroblasts and retinal epithelial cells. Further, this reorganization alters the accessibility of predicted binding sites for the TEAD TF family, and avoids closing the binding sites of CTCF, in both cell types.

TEAD1 transcription factor genomic binding is substantially depleted upon HCMV infection

Because chromatin re-organization was more pronounced in fibroblasts (**Figure 2**), we focused on this cell type for subsequent chromatin immunoprecipitation experiments and analyses. We prioritized the two TF families (TEAD and CTCF) with consistent enrichment patterns across cell types within HCMV-dependent chromatin accessibility. We selected TEAD1 to represent the TEAD family because it is the only member whose expression levels are significantly altered by HCMV infection (see **Figure 5D**). We also examined the genome-wide distribution of H3K27ac, a histone mark correlated with active enhancers.

ChIP-seq experiments for TEAD1, CTCF, and H3K27ac were performed in duplicate in uninfected and HCMV infected conditions (48 hours post infection). The resulting data were of high quality, meeting or exceeding ENCODE data quality standards (**Supplemental Dataset 7**), with strong agreement between replicates (**Supplemental Figure S2**). In particular, all H3K27ac datasets had > 65,000 peaks and all TEAD1 and CTCF ChIP-seq datasets had > 15,000 peaks, with FRiP scores ranging from 0.02 to 0.13 and very strong motif enrichment results (CTCF and TEAD motifs rank #1 in every respective experiment, with p-values < 10^{-3000})

(**Supplemental Dataset 7**). As predicted, the number of TEAD1 ChIP-seq peaks was substantially diminished in infected cells (75,554 in uninfected compared to 17,567 peaks in infected cells), along with a substantial drop in H3K27ac peaks (110,308 vs. 66,644) (**Supplemental Dataset 7**), suggesting that loss of TEAD1 binding might have a strong impact on enhancer functionality. In contrast, the total number of CTCF peaks was largely consistent with and without infection (54,697 vs. 57,936 peaks).

To further quantify these changes, we performed differential ChIP-seq peak analysis using the DiffBind software package (Ross-Innes et al., 2012) (see Methods). For TEAD1, we identified 30,740 peaks specific to the uninfected condition, with only 237 peaks specific to infected cells. Likewise, we identified 40,656 H3K27ac peaks specific to uninfected cells, with 8,542 peaks specific to infected cells (**Supplemental Dataset 8**). As predicted by the motif enrichment analysis of differentially accessible chromatin, regions where the chromatin closed with infection did not contain CTCF ChIP-seq peaks. Instead, we found that losses of TEAD1 occupancy strongly coincide with losses in chromatin accessibility (enrichment: 9.1-fold, adjusted p-value: 1.4×10^{-215}), losses in chromatin looping events (enrichment: 2.1-fold, adjusted p-value: 3.5×10^{-192}), and losses in H3K27ac levels (enrichment: 6.1-fold, adjusted p-value: 8.7×10^{-214}) (**Figure 4A** and **Supplemental Dataset 5**). For example, we observe significant HCMV-induced loss of TEAD1 binding, chromatin accessibility, and H3K27ac levels proximal to the promoters of the Hippo pathway genes *FRMD6* and *RASSF2* (**Figure 4B**), both of which have significantly diminished gene expression subsequent to infection (see next section). In the case of *FRDM6*, which involves a likely enhancer, we also observe loss of a chromatin looping interaction (blue loop) between the promoter (dashed box) and two TEAD1 binding sites lost upon infection (solid box). In the case of *RASSF2*, a TEAD1 binding site within the promoter (solid box) interacts (blue loop) with a downstream enhancer (dashed box). Both the TEAD1 binding site and the interaction are lost upon infection. These examples indicate that loss of TEAD1 binding rather than CTCF binding correlates with loss of chromatin interactions. Collectively, these observations confirm our motif-based predictions that loss of chromatin accessibility co-occurs with loss of TEAD1 binding, in addition to loss of enhancer function (H3K27ac) and enhancer looping interactions (HiChIP).

HCMV infection induces large-scale changes to human gene expression, including perturbation of the Hippo signaling pathway

Our data indicate substantial changes to human gene regulatory features upon HCMV infection. Therefore, we next systematically assessed HCMV-dependent alterations to human gene expression levels. To this end, RNA was extracted from uninfected and infected human fibroblasts at 48 hours post infection and analyzed by RNA-seq (see Methods). The resulting data were of high quality (**Supplemental Dataset 9**) and displayed strong agreement between replicates (**Supplemental Figure S3**). As expected, we detected HCMV-encoded genes only in infected cells (**Supplemental Dataset 10**). We identified 2,352 differentially regulated human genes (1,155 upregulated with infection and 1,197 downregulated) at a 2-fold cutoff with an FDR < 0.01 (**Figure 5A** and **Supplemental Dataset 11**).

As expected, pathway enrichment analysis of differentially expressed genes revealed strong enrichment for pathways generally involved in the infection response, including “viral protein interaction with cytokine and cytokine receptor” and “human papillomavirus infection” (**Figure 5B**). We also observe strong enrichment for many of the major developmental pathways, including WNT, TGF-beta, and Hippo (**Figure 5B**). Strikingly, many of the genomic regions with extensive (five or more) TEAD1 binding loss events encode major regulators of these pathways, including *WNT5B* and *FZD7* (WNT pathway), *SMAD3*, *TGFB2*, and *TGFBR2* (TGF-beta pathway), and *TEAD1* itself (Hippo pathway), all of which have HCMV-altered expression levels (**Table 1**, **Supplemental Dataset 12**). The extensive loss of TEAD1 binding (with concordant alterations to gene expression levels) suggests that TEAD1 might be a key HCMV-targeted regulator. TEAD transcription factors are direct effectors of the Hippo signaling pathway (Currey et al., 2021), and our data support TEAD being at the top of a hierarchy involving many other key developmental pathways.

Overall, we observe extensive HCMV-induced alterations to Hippo gene expression levels. Of the 163 Hippo pathway genes annotated in the KEGG database, 21 were significantly downregulated (including *TEAD1*) and 20 were upregulated (including *PARD6B*, a negative regulator of Hippo (Frum et al., 2018)) (**Supplemental Dataset 13** and **Figure 5C**). Notably, only TEAD1 was differentially expressed among the four TEAD family members, with a 3.7-fold decrease upon infection (adjusted p-vale: 1.9×10^{-29}) (**Figure 5D** and **Supplemental Dataset 11**). We confirmed alterations in expression levels at the protein level for the classic TEAD1 targets Thrombospondin 1 (THBS1) and Cellular Communication Network Factor 1 (CCN1), both of which were substantially downregulated upon HCMV infection. Specifically, THBS1 protein was not detected in infected cells, while CCN1 was reduced by over 3-fold (**Figure 5E**).

HCMV infection diminishes TEAD1 transcription factor activity through four distinct mechanisms

We next aimed to further explore the virus-induced mechanisms underlying the reduction in TEAD1 activity. The transcriptional regulatory activity of TEAD1 is primarily controlled by the phosphorylation-dephosphorylation cycle of its co-activator YAP1 (Ma et al., 2019; Totaro et al., 2018). We thus examined the protein expression levels and phosphorylation status of YAP1 (pYAP1) with and without infection, along with the expression of HCMV proteins IE1/2, TEAD1, and the H3K27ac histone mark in uninfected and infected cells (**Figure 6A**, **Supplemental Figure S4**). Consistent with diminished TEAD1 activity, we observed a reduction in levels of total YAP1 protein (1.8-fold, $p<0.03$) and pYAP1 levels (1.5-fold, $p<0.0005$) (**Figure 6A, Supplemental Figure S4**). It is likely that the reduction of pYAP1 is a direct function of reduced YAP1 expression. These results are consistent with a large-scale HCMV infection proteomics screen performed in human fetal foreskin fibroblasts (Weekes et al., 2014). Likewise, TEAD1 protein levels decreased (3.7-fold, $p<0.0009$) with HCMV infection (**Figure 6A, Supplemental Figure S4**). Consistent with the H3K27ac ChIP-seq results (**Figure 4B**), global H3K27ac was also substantially reduced (1.9-fold, $p<0.19$) in infected cells. Since YAP1 activity is primarily controlled by phosphorylation of its serine residues (S127 and S381), which leads to its cytoplasmic translocation (Pocaterra et al., 2020), we next assessed the cytoplasmic-nuclear shuttling of pYAP1. These experiments revealed increased YAP1 and pYAP1 in the nuclear fraction of uninfected cells relative to HCMV infected cells (**Figure 6B**). Taken together, these results indicate that HCMV-induced loss of TEAD1 activity can also partially be accounted for by loss of active YAP1.

Recent reports have shown that alternative splicing of *TEAD1* also regulates TEAD1 activity (Choi et al., 2022). We thus used the AltAnalyze software package (Emig et al., 2010) to systematically examine alternative splicing changes between HCMV infected and uninfected cells within our RNA-seq data (see Methods). This analysis identified two differential exon use events within the *TEAD1* gene that are impacted by HCMV infection: 1) skipping of exon 6 with HCMV infection; and 2) diminished inclusion of an upstream intronic region just upstream (~90 bases) of exon 5 with HCMV infection (**Supplemental Figure S5**). A recent study carefully examined the functional effect of the alternative exon 6 usage event, concluding that skipping of exon 6 diminishes TEAD1 activity by disrupting the intramolecular interaction between its DNA-binding domain and the YAP1-interacting domain (Choi et al., 2022). We confirmed HCMV-

induced differential splicing of TEAD1 exon 6 by RT-PCR of total RNA isolated from uninfected and HCMV infected fibroblasts (**Figure 6C** and **Supplemental Figure S5C**).

Taken together, our results reveal multiple interrelated mechanisms by which HCMV infection impairs TEAD1 activity (**Figure 6D**): 1) Exclusion of TEAD1 exon 6 (**Figure 6C**); 2) Reduction of *TEAD1* gene and TEAD1 protein levels (**Figures 5D** and **6A**); 3) Reduction of YAP1 and pYAP1 protein levels (**Figure 6A**); and 4) Extensive closing of human chromatin that is normally occupied by TEAD1 (**Figure 4**).

HCMV-induced TEAD1 binding loss coincides with genetic variants associated with ear and eye growth defect phenotypes

Next, we examined the enrichment of GWAS hits within regions of HCMV-induced loss of TEAD1 binding. To this end, we used our RELI tool (Harley et al., 2018) to estimate GWAS-associated risk locus enrichment for 1337 diseases and phenotypes. This analysis produced three phenotypes with specific enrichment for TEAD1 binding loss events compared to unchanged TEAD1 binding events (see Methods and **Supplemental Dataset 14**): lobe attachment (Adhikari et al., 2015; Shaffer et al., 2017), optic cup area (Bonnemaijer et al., 2019; Gharahkhani et al., 2018; Springelkamp et al., 2017; Springelkamp et al., 2015), and percent mammographic density (Lindstrom et al., 2014; Lindstrom et al., 2011; Sieh et al., 2020). Among these, enrichment replicated in multiple GWAS for two phenotypes: lobe attachment and optic cup area (**Figure 6E**). Notably, HCMV infection of fibroblasts has previously been shown to contribute to hearing impairment (Fowler and Boppana, 2006; Goderis et al., 2014; Nystad et al., 2008; Singh et al., 2023b) and MCMV infection can lead to hearing loss in mice (Bradford et al., 2015). Likewise, Hippo signaling plays established roles in ear growth and hearing loss (Deng et al., 2016; Gnedeva et al., 2020; Holden and Cunningham, 2018; Wang et al., 2022). Of note, the thrombospondin (*THBS1*) gene, which plays important, well-established roles in hearing loss (Smeriglio et al., 2019), has extensive proximal HCMV-induced TEAD1 binding loss and HCMV-induced gene (**Supplemental Dataset 11**) and protein level reduction (**Figure 5E**). Meanwhile, the “optic cup area” phenotype refers to the overall size of the optic nerve, a phenotype influenced by HCMV infection of the eye (Ijezie et al., 2023; Xu et al., 2020). Hippo signaling plays well-established roles in eye development (Moon and Kim, 2018). Collectively, these results implicate HCMV-induced TEAD1 binding loss in both hearing and eyesight growth

defects at multiple specific genomic loci, providing mechanistic insights into the well-established roles of both HCMV and Hippo signaling in these disorders.

Discussion

HCMV is a leading cause of birth defects in developing fetuses, including ear and eye developmental disorders. Upon infection, HCMV alters the expression of thousands of human genes. A better understanding of the mechanisms underlying these changes is critical for understanding the processes underlying HCMV-associated diseases. In this study, we used several complementary functional genomics approaches along with biochemical validation to show that HCMV employs multiple interrelated molecular mechanisms that diminish the activity of the human TEAD1 transcription factor: closing of chromatin in regions normally bound by TEAD1 (with concomitant reduction of H3K27ac levels and chromatin looping interactions), lowering of TEAD1 protein expression levels, preferential exclusion of TEAD1 exon 6, and lowering of protein expression of the TEAD1 co-activator YAP1. This widespread reduction in TEAD1 activity results in substantial alteration to the expression of many human genes, including targets of Hippo and other key developmental signaling pathways.

The human genome encodes four TEAD transcription factors (TEAD1-4), all of which are mediators of the Hippo signaling pathway. The Hippo pathway regulates key aspects of development, including control of organ size, stem cell identity, and lineage specification (Currey et al., 2021; Ma et al., 2019). The activity of TEAD transcription factors (TFs) is regulated by the availability of their co-factors YAP1, TAZ (WWTR1), and members of the vestigial-like protein family (VGLL1-4). While YAP1 and TAZ function as transcriptional co-activators of TEAD TFs, VGLL proteins function as co-repressors (Currey et al., 2021; Rausch and Hansen, 2020; Yamaguchi, 2020). The activity of TEAD1 has previously been shown to be affected by differential use of exon 6, with exclusion of exon 6 reducing activity due to an impaired ability to interact with YAP1 (Choi et al., 2022). The YAP1-TEAD complex exerts its activity by binding to enhancers of genes involved in extracellular matrix organization, actin cytoskeleton organization, and cell adhesion (Rausch and Hansen, 2020; Stein et al., 2015), all of which are pathways for which we see significantly altered gene expression levels subsequent to HCMV infection (**Figure 5B**).

A previous report has shown that HCMV infection inhibits the proliferation and invasion of extravillous cytotrophoblasts (EVT) through mRNA downregulation of core Hippo pathway

components, including downregulation of YAP1/TAZ and all four TEAD TFs (Kong et al., 2021). Here, we observe a 3-fold depletion in TEAD1 protein levels and a substantial reduction in YAP1 protein levels, in addition to other mechanisms impacting TEAD1 activity. Additionally, the expression of key TEAD1 targets, including Connective Tissue Growth Factor (CTGF), Fibronectin (FN), α-smooth muscle actin (ACTA2) and Collagen 1a (COL1A) have all previously been found to be downregulated with HCMV infection of trabecular meshwork cells, leading to an increase in intra-ocular pressure, a leading cause of glaucoma (Choi et al., 2017). Although the exact mechanisms seem to differ depending on the cell type, HCMV infection of permissible cells clearly perturbs the Hippo pathway in multiple cellular contexts by diminishing TEAD TF activity.

Recent studies have shown that other viruses perturb the Hippo pathway as part of their replicative cycle (Wang et al., 2019). For example, oncogenic viruses, such as Kaposi's sarcoma-associated herpesvirus (KSHV) and Human papillomavirus (HPV) promote dephosphorylation and nuclear translocation of YAP1 (Liu et al., 2015). In the case of KSHV, the viral GPCR (vGPCR; ORF74) has been shown to inhibit the Hippo kinases LATS1/2, thereby activating YAP1/TAZ (Liu et al., 2015). Likewise, HPV16 oncoprotein E6 induces cell proliferation by dephosphorylating a serine residue of YAP1 (S379) and preventing its degradation by the proteasome complex (He et al., 2015). More recently, it has been discovered that Epstein-Barr virus subverts the YAP/TAZ pathway in lytic reactivation in epithelial cells (Singh et al., 2023a; Van Sciver et al., 2021).

The results of this study represent one of likely many examples of a single virus targeting the same host molecule through multiple mechanisms. For example, Human Papillomavirus targets TP53 through independent mechanisms involving both the E6 and E7 proteins (Fontan et al., 2022). Likewise, poxviruses target caspase activity through multiple mechanisms (Nichols et al., 2017). Given the vast amount of evolutionary time that viruses have at their disposal due to their highly elevated mutation rates, and the reliance of viruses on host-encoded pathways, it is likely that many other viruses also target a single molecule through multiple mechanisms.

Several previous studies have examined changes to human gene expression induced by HCMV infection in a variety of contexts (reviewed in (Marti-Carreras and Maes, 2019)). To date, only two studies have employed more than one type of genome-scale measurement to compare HCMV infected to uninfected cells. In the first study, multiple single cell approaches were employed (including scATAC-seq and CITE-seq) to compare NK cells with or without infection, revealing a possible role for AP-1 family TFs (Ruckert et al., 2022). In the second study,

integration of Pol2, H3K27ac, and H3K27me3 ChIP-seq data with ATAC-seq data revealed epigenetic reprogramming of the virus and host genomes in myeloid progenitor cells (Forte et al., 2021). Interestingly, this study found only limited changes to chromatin accessibility between infected and uninfected cells, likely due to differences in cell type (Kasumi-3 vs. HS-68 and ARPE-19 in our study), time point (24 vs. 48 hours), infection strength (MOI of 1 with <50% infection vs MOI of 5 and 10 with >90% infection), and replication efficiency (~30 vs. ~250 viral genomes per cell). To our knowledge, our study is the first to employ a systematic approach to identify important host transcriptional regulators through unbiased analysis of virus-altered chromatin accessibility regions, with subsequent ChIP-seq-based validation. Future studies will be needed to systematically identify important host-encoded regulators in the context of other cell types and other viruses.

HCMV infects up to 86% of the world's population (Zuhair et al., 2019). For most, this infection is largely benign. However, infection can cause complications in newborn infants and the immunocompromised. Our comprehensive analyses reveal that HCMV targets the Hippo signaling pathway through TEAD1, setting off a cascade of effects on other key developmental pathways. Importantly, these TEAD1 binding loss events are highly enriched for genetic variants associated with eye and ear phenotypes, providing possible mechanistic insights into the well-established role of HCMV infection in eye and ear disorders. Collectively, the results of our study offers new avenues to investigate HCMV-dependent mechanisms in healthy and disease states.

Methods

Experimental design, cell culture, and viral infections

Experiments were performed in human foreskin fibroblasts (HS68 cells). ATAC-seq experiments were also performed in human retinal epithelial cells (ARPE-19 cells). HCMV infections were performed using the TB40/E clinical isolate. Fibroblasts were infected at a multiplicity of infection (MOI) of 5 and cells were harvested 48 hours post infection for RNA-seq, ATAC-seq, ChIP-seq, and HiChIP (**Figure 1**). Retinal epithelial cells were infected at an MOI of 10. The rate of infection was monitored by both FACS and microscopy.

Gene expression (RNA-seq)

Total RNA was isolated from $5x10^6$ cells, either uninfected or infected (48 hrs), using the RNeasy RNA Isolation kit (Qiagen# 74104) as per instructions provided by the manufacturer. A total of 500 ng of RNA per sample was used as input material for ribosomal-RNA depleted RNA-sequencing. First, ribosomal RNA (rRNA) was removed, and rRNA free residue was cleaned up by ethanol precipitation. Subsequently, sequencing libraries were generated using the rRNA-depleted RNA by the Directional RNA Library Prep Kit. Briefly, fragmentation was carried out using divalent cations under elevated temperature in first strand synthesis reaction buffer (5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNaseH-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. In the reaction buffer, dNTPs with dTTP were replaced by dUTP. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, adaptors with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with the AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 µL USER Enzyme (NEB, Ipswich, MA, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 minutes followed by 5 minutes at 95 °C before PCR. PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index Primer. Finally, products were purified (AMPure XP system) and library quality was assessed on an Agilent Bioanalyzer 2100 system (Agilent Technologies, Inc; Santa Clara, CA, USA). The libraries were sequenced on an Illumina NovaSeq 6000 at Novogene (paired-end, 150 bp read length).

RNA sequencing data were processed using the nf-core/rnaseq pipeline (version 3.8.1) (Di Tommaso et al., 2017; Ewels et al., 2020; Harshil et al., 2022). Initial quality control of the raw sequencing data was performed using FastQC (version 0.11.9) (Andrews, 2010). Low-quality bases and adapter sequences were then trimmed and filtered from the reads using Cutadapt (version 3.4) (Martin, 2011) and Trim Galore (version 0.6.7) (Krueger et al., 2021), respectively. Ribosomal RNA sequences were subsequently removed from the aligned data using SortMeRNA (version 4.3.4) (Kopylova et al., 2012) to eliminate potential contamination from non-target RNA species. STAR aligner (version 2.7.10a) (Dobin et al., 2013) was used to align the trimmed reads to a custom reference genome of hg19 and Human herpesvirus 5 strain TB40/E clone TB40-BAC4 sequence (GenBank: EF999921.1). The resulting alignments were sorted and indexed using SAMtools (version 1.15.1). To further evaluate the quality of the RNA sequencing data, several tools were employed, including RseqQC (version 3.0.1) (Wang et al., 2012), Qualimap (version 2.2.2-dev) (Okonechnikov et al., 2016), dupRadar (version 1.18.0)

(Sayols et al., 2016), and Preseq (version 3.1.1)(Daley and Smith, 2013). These tools provide comprehensive assessments of various quality metrics, such as read distribution, GC content, duplication rates, and library complexity, ensuring reliable data for downstream analysis. For transcript quantification, Salmon (version 1.5.2) (Patro et al., 2017) was utilized to estimate the expression levels of transcripts. Finally, differential gene expression analysis was performed using DESeq2 (version 1.30.1) (Love et al., 2014). Genes were considered differentially expressed if they had a 2-fold change and an adjusted p-value threshold of less than 0.01. Pathway enrichment analyses for the differentially expressed genes were performed using Enrichr (Chen et al., 2013; Kuleshov et al., 2016). An adjusted p-value threshold of 0.05 was used to generate the pathway enrichment figure (**Figure 5B**).

Alternative splicing analysis

Exon-exon and exon-intron spanning reads were identified with the software AltAnalyze (version 2.1.4) using the Ensembl-72 human database, along with splicing event calculation and annotation with the MultiPath-PSI algorithm (see: http://altanalyze.readthedocs.io/en/latest/Algorithms for algorithm details and benchmarking). To detect high-confidence splicing changes between HCMV infected cells and uninfected cells, splicing events that were not detected in all replicates in each biological group were excluded from the differential splicing analysis. Significant splicing changes were defined as splicing events with a change in percent spliced in (PSI) value between the two groups > 10% (ΔPSI > |0.1|), with p-value <0.01.

Chromatin accessibility (ATAC-seq)

Omni-ATAC-Seq was performed in replicates as previously published (Corces et al., 2017). Briefly, HCMV infected and uninfected control cells were harvested by trypsinization for 5 minutes. Approximately 70,000 cells were transferred to a microfuge tube and washed once with PBS. The cell pellet was resuspended in 50 µL of ice-cold lysis buffer (10 mM Tris-HCl [pH 7.5], 10 mM NaCl, 3 mM MgCl2, 0.1% NP-40, 0.1% Tween-20, and 0.01% Digitonin) and incubated on ice for 3 minutes, followed by centrifugation at 500g for 5 minutes at 4 °C. For the transposition reaction, nuclei were resuspended in 50 µL of Nextera transposition reaction mix consisting of 25 µL 2x TD Buffer, 2.5 µL Nextera Tn5 transposase (Illumina # FC-121-1030), and 22.5 µL of nuclease free water. The reaction mixture was mixed 5-10 times by gentle pipetting and incubated at 37 °C for 45 minutes. The transposed DNA was then purified by Qiagen MinElute kit (Qiagen # 28004) and eluted in 11 µL of elution buffer. 1 µL of the eluted

DNA was used to assess quality on an Agilent Tapestation 4510. For library preparation, 10 µL of tagmented DNA was PCR amplified for 14 cycles in a 50 µL reaction volume using NEBNext High-Fidelity 2X PCR Master Mix (NEB # M0541S) and Nextera primers. The amplified library was purified and size-selected using AMPure XP beads (Beckman # 63880) in a two-step protocol. First, DNA fragments >1000 bp were removed from the PCR reaction mixture by adding 22.5 µL (0.5x volume) of AMPure XP beads followed by incubation at room temperature (RT) for 10 minutes. The beads were magnetized, and the supernatant was transferred to a new tube. In the second size selection step, 58.5 µL (1.2x volume) of AMPure XP beads were added to the supernatant, mixed by pipetting, and incubated for 10 minutes at room temperature. Beads were washed twice with 80% ethanol, air-dried and DNA eluted in 30 µL of 10 mM Tris-HCl [pH 8.0]. Library DNA concentrations and quality were assessed by a Qubit dsDNA HS assay kit (Q32854) and Agilent 4510 TapeStation system, respectively. Libraries were sequenced on an Illumina NovaSeq 6000 at the Cincinnati Children's Hospital Medical Center Genomics Sequencing Facility (paired-end, 150 bp read length).

ATAC-seq data were processed and aligned to the hg19 genome using the ENCODE ATAC-seq pipeline (V2.0.0) (Consortium, 2012; Hitz et al., 2023; Lee et al., 2021a; Luo et al., 2020). Peaks were called within the pipeline using MACS2 (Zhang et al., 2008). Differential chromatin accessibility analysis was performed using DiffBind 3.6.5 (Rory Stark, 2017; Ross-Innes et al., 2012) in R 4.2.1 (Team, 2023). The "conservative overlap" peak sets generated from the individual replicates by the ENCODE ATAC-seq pipeline were used for differential analysis. Peaks were considered to be differentially accessible if the FDR was less than 0.01 and the fold change was 2-fold or greater. A modified version of HOMER (Heinz et al., 2010) using a log base 2 likelihood scoring system was used to calculate motif enrichment statistics for a large library of human position weight matrix (PWM) binding site models contained in build 2.0 of the CisBP database (Weirauch et al., 2014). DeepTools (v2.0.0) (Ramirez et al., 2016) was used to generate heatmaps of signal tracks across differentially accessible chromatin for each set of comparisons.

Chromatin immunoprecipitation sequencing (ChIP-seq)

2 x 10^7 cells were seeded in 150 mm culture dishes and were either infected with HCMV at an MOI of 5 or kept uninfected. After 48 hours, the medium was removed, and cells were harvested by trypsinization followed by two washes with PBS. Cells were crosslinked with 10 mL of formaldehyde cross-linking solution (1x PBS, and 1% formaldehyde) for 10 minutes at RT. Crosslinking reaction was quenched by adding 2.5 M Glycine to a final concentration of 125 mM

and incubated for 5 minutes at RT. Cells were washed twice with ice-cold 1x PBS. Crosslinked cells were then transferred to a microfuge tube and frozen at −80 °C until further use.

For chromatin preparation, crosslinked cells were resuspended in L1 buffer (50 mM Hepes-KOH [pH 8.0], 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, and 1x protease inhibitors) and incubated at 4 °C on a rotator for 10 minutes. Nuclei were pelleted by centrifugation and resuspended in L2 buffer (10 mM Tris-HCl [pH 8.0], 200 mM NaCl, 0.5 mM EGTA, and 1x protease inhibitors) with rotation for 10 minutes at RT. Isolated nuclei were then resuspended in sonication buffer solution (10 mM Tris-HCl, 1 mM EDTA, and 0.1% SDS). To obtain chromatin fragments of 200-500 bp length, nuclei were sonicated using an S220 ultrasonicator (Covaris, LLC., Woburn, MA, USA) at 10% duty cycle, 175 peak power, 200 burst/cycle for 7 minutes at 4 °C. A portion of the sonicated chromatin was run on an agarose gel to verify fragment sizes. The chromatin solution was centrifuged to pellet the debris and the supernatant was collected in a fresh microfuge tube. Since the sonicated chromatin solution is devoid of any detergents or salts, the following were added at indicated final concentrations: Triton X-100 (1%), Sodium deoxycholate (0.1%), Glycerol (5%), NaCl (150 Mm), and Protease inhibitor cocktail (1x). Chromatin was pre-cleared with Protein G Dynabeads for 45 minutes at 4 °C on a rotator. Chromatin immunoprecipitations were performed in an SX-8X IP-Star Compact automation system (Diagenode. LLC.,Denville, NJ CA, USA) with 200 µL of pre-cleared chromatin (approximately 3 million cells), 21 µL of Protein A or G Dynabeads (Thermo Fisher Scientific, Waltham, MA USA), and 5 µg of antibody. The following antibodies were used for ChIP: CTCF (ActiveMotif # 61311, lot 17118005), TEAD1 (ActiveMotif # 61643, lot 34614001) and H3K27ac (Abcam # ab4729, lot GR3357415-1). Immunoprecipitation of chromatin was carried-out for 8 hours, after which the Dynabeads were sequentially washed with Wash Buffer 1 (10 mM Tris-HCl [pH 8.0], 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, and 1% Triton X-100), Wash Buffer 2 (10 mM Tris-HCl [pH 8.0], 250 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, and 1% Triton X-100), Wash Buffer 3 (50 mM Tris-HCl [pH 8.0], 2 mM EDTA, and 0.2% N-Lauroylsarcosine sodium salt), and Wash Buffer 4 (TE + 0.2% Triton X-100) for 5 minutes each. For library preparation, immunoprecipitated chromatin was eluted in elution buffer (1x TE, 250 mM NaCl and 0.3% SDS). Chromatin proteins and RNA were digested with Proteinase-K and RNase A respectively and DNA purified with the Qiagen MinElute kit. A ChIP-DNA library was constructed using the NEBNext Ultra II DNA library preparation kit (E7645S) as per manufacturer's instructions and purified with 0.6x volume of AMPure XP beads. Purified library DNA quality and quantity were assessed by an Agilent TapeStation 4150 and sequenced using an Illumina NovaSeq 6000 at

the Cincinnati Children's Hospital Medical Center Genomics Sequencing Facility (Single end, 100 bp read length).

ChIP-seq data were processed and aligned to the hg19 genome using the ENCODE ChIP-seq pipeline (V2.0.0) (Consortium, 2012; Hitz et al., 2023; Lee et al., 2021b; Luo et al., 2020). Peaks were called within the pipeline using MACS2 (Zhang et al., 2008). Differential peak analysis was performed using DiffBind 3.6.5 (Rory Stark, 2017; Ross-Innes et al., 2012) in R 4.2.1 (Team, 2023). The "conservative overlap" peak sets generated from the individual replicates by the ENCODE ChIP-seq pipeline were used for differential analysis. Peaks were considered differentially enriched if the FDR was less than 0.01 and the fold change was 2-fold or greater.

Chromatin looping interactions (HiChIP)

HiChIP libraries were prepared following the protocol from Mumbach et al (Mumbach et al., 2016). Experiments were performed using biological triplicates in both uninfected and infected cells. In brief, 10 million cells were cross-linked with 1% formaldehyde, followed by a cell lysis with Hi-C lysis buffer (10 mM Tris-HCl [pH 8.0], 10 mM NaCl, 0.2% NP-40, 1x Roche protease inhibitors) for 30 minutes at 4 °C. Following cell lysis, 375 U of MboI (NEB, R0147) was used to digest, *in-situ*, the cross-linked chromatin for 2 hours at 37 °C. After filling-in the DNA ends by biotin-dTAP (Thermo # 19524016), dCTP, dGTP, and dTTP with 5 U/µL DNA polymerase (NEB # M0210), the DNA was ligated by T4 DNA ligase (NEB # M0202) at 4 °C overnight. The ligated DNA was sonicated using a Covaris S220 platform (Covaris, LLC., Woburn, MA, USA) in nuclear lysis buffer (50 mM Tris-HCl [pH 7.5], 10 mM NaCl, 0.2% NP-40, 1x Roche protease inhibitors) at 4 °C. The fragmented DNA was then diluted 10 times with ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl [pH 7.5], 167 mM NaCl), and the samples were precleared with Protein A Dynabeads at 4 °C for 1 hour, followed by immunoprecipitation with 8 µL H3K27Ac antibody (Abcam # ab4729, lot GR3374555-1) at 4 °C overnight. DNA-protein complexes were captured by Protein A beads with rotation at 4 °C for 2 hours. The protein A beads were then washed sequentially with low salt, high salt, and LiCl wash buffer, followed by elution twice with 100 µL freshly prepared DNA elution buffer at 37 °C (50 mM $NaHCO_3$, 1% SDS). The eluted chromatin was reverse cross-linked and purified by a PCR purification Kit (Qiagen). 5 µL of Streptavidin C-1 beads (Thermo Fisher # 65001) was used to capture the Biotin Datp labeled DNA. The captured DNA was transposed with 2.5 µL Tn5 transposase (Illumina). The beads were then sequentially washed by Tween wash buffer (5 mM Tris-HCl [pH 7.5], 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20), 50 mM EDTA, and 10 mM Tris. After washing, PCR was performed by re-suspending the beads with 23 µL H_2O, 25 µL 2X

Phusion HF (New England Biosciences), 1 µL Nextera forward primer (Ad1_noMX), and 1 µL Nextera reverse primer (Ad2.X) at 12.5 µM. The PCR was run at 8 cycles of (1) 72 °C for 5 minutes, (2) 98 °C for 1 minute, (3) 98 °C for 15 seconds, (4) 63 °C for 30 seconds, followed by an extension at 72 °C for 1 minute. The post-PCR size selection was performed by two-sided selection with AMPure XP beads to capture the fragments between 300 and 700 bp. Samples were then sequenced on an Illumina NovaSeq 6000 at Novogene (paired-end, 150 bp read length).

HiC-Pro (version 2.11.4) was used to align and filter read pairs and identify the contact map (Servant et al., 2015). MboI restriction sites and default parameters were used to align the reads to the hg19 genome. After alignment, the read pairs were filtered to remove those that mapped to multiple locations, were not in a valid orientation, or were duplicated. Quality control metrics were generated at each step. HiC-Pro also reports the number of *trans* pairs, short-range *cis* pairs and long-range *cis* pairs, as well as the number of pairs in each orientation (**Supplemental Dataset 4**). The overlap conservative peaks generated by the ENCODE pipeline for the H3K27ac ChIP-seq data were used as input for calling loops using FitHiChIP (version 11.0) (Bhattacharyya et al., 2019). The parameters for FitHiChIP were the following: interaction type of “peak to all”; bin size of 10 kb; lower loop distance of 20 kb; upper loop distance of 20 mb; background model of loose; bias correction of coverage bias regression; merge filtering enabled; and FDR of 0.01 (**Supplemental Dataset 3**). Peaks were called from the HiChIP data using HiChIP-Peaks (version 0.1.2) (Shi et al., 2020) using the MboI restriction sites and an FDR of 0.01. The peaks were compared to the H3K27ac ChIP-seq peaks using bedtools (version 2.30.0) (Quinlan and Hall, 2010) (**Supplemental Dataset 4**). We observed strong agreement between experimental replicates (**Supplemental Figure S6**), so we pooled the reads for the replicates and analyzed the resulting data as described above (**Supplemental Dataset 3 and 4**). We identified differential looping events using the DiffAnalysisHiChIP.r script included in the FitHiChIP software with a default bcv parameter value of 0.4 and cutoffs of >2-fold change and FDR < 0.01 (**Supplemental Dataset 3**). We used the merged loops from the combined replicates as input. Replicate comparisons (**Supplemental Figure 6**) were performed using the multiBamSummary and plotCorrelation programs from the deepTools software package (Ramirez et al., 2016). The parameters used for multiBamSummary were --binSize 10000 and --distanceBetweenBins 0. The parameters used for plotCorrelation were --corMethod pearson, --skipZeros, and -- removeOutliers.

Assessment of protein expression levels (Western blots)

Uninfected and HCMV infected cells were harvested, in triplicates, by trypsinization and washed twice with PBS. The cell pellet was then resuspended in 1x RIPA buffer supplemented with Halt Protease inhibitor cocktail (Sigma) and kept on ice for 30 minutes. Cell lysates were cleared by centrifugation at 12,000 RPM for 10 minutes and supernatant was collected in a new tube. Protein concentration was measured by the BCA method and 30 µg of protein loaded on 4-12% Bis-tris novex gel (Invitrogen). After electrophoresis, proteins were transferred on to a Nylon membrane using Invitrogen's semi-dry transfer method. Proteins were blocked with Intercept blocking buffer (LI-COR) and membrane incubated with HCMV IE1/2 (Millipore # MAB810R), TEAD1 (Cell Signaling # 12292), YAP1 (Cell Signaling # 14074), pYAP1 (Phospho-YAP1; Ser127; Cell Signaling # 4911), THBS1 (Novus Biologicals # NB100-2059), CCN1 (Novus Biologicals # NB100-356), H3K27ac (Abcam # ab4729), β-Actin (Cell Signaling # 8457) or GAPDH (Invitrogen # ma5-15738) antibodies at 4 °C for overnight. Following antibody binding, membrane was washed 3 times with 1x TBST and incubated either with IRDye® 800CW or IRDye® 680RD conjugated secondary antibodies for 45 minutes at RT. For target protein quantification against a loading control (GAPDH and β-Actin), secondary antibodies against both target protein and loading control were added to the blot. The membrane was washed 3 times with 1x TBST and imaged on an Odyssey Dlx system (LI-COR, Biosciences LLC., Lincoln, NE, U.S.A). Scanned images were then used for target protein quantification using Emperia Studio software (LI-COR, Biosciences LLC., Lincoln, NE, U.S.A).

Assessment of YAP1 localization

Uninfected and HCMV infected cells were harvested by trypsinization and washed twice with PBS. Cytosolic and nuclear protein extracts were prepared using Pierce NE-PER Nuclear and Cytoplasmic Extraction Reagent Kit (Thermo Scientific # 78833), as per manufacturer's instructions. Protein concentration was measured by the BCA method and 15 µg of protein loaded on 4-12% Bis-tris novex gel (Thermo Fisher Scientific, Waltham, MA USA). Western blotting and imaging were performed as given above with the following antibodies: YAP1 (Cell Signaling # 14074), pYAP1 (Phospho-YAP1; Ser127; Cell Signaling # 4911), Histone H3 (Abcam # ab1791) and GAPDH (Invitrogen # ma5-15738).

Validation of TEAD1 splicing (RT-PCR)

Total RNA was extracted from uninfected and HCMV infected fibroblasts 48 hours post infection (hpi) using the RNA mini kit from Qiagen as per manufacturer's instructions. 2.5 µg of total RNA was reverse transcribed using SuperScript™ IV VILO™ Master Mix (Thermo Fisher Scientific,

Waltham, MA USA) and 3 µL of cDNA was used in PCR to amplify the regions immediately flanking TEAD1 exon-6 using forward and reverse primers. Forward primer: 5'-ATCTCGTGATTTTCATTCCAAGC. Reverse Primer: 5'-TGAGGACATGGCCGCCATGTGC.

Data Visualization

Visualization tracks for each dataset were created for the UCSC Genome Browser (Kent et al., 2002). Signal tracks (in bigWig format) were created using the bamCoverage program from the deepTools software package (Ramirez et al., 2016) with the parameters --normalizeUsing BPM and --binSize 10. For the RNA-seq signal tracks, the command was run twice in order to maintain strandedness. For the forward strand track, the parameter --filterRNAstrand forward was included. For the reverse strand track, the parameters --filterRNAstrand reverse and --scaleFactor -1 were included. The .hic files for the HiChIP data were created using the hicpro2juicebox.sh script from HiC-Pro (version 2.11.4) (Servant et al., 2015). Juicer (version 1.22.01) (Durand et al., 2016) was used for KR normalization.

TEAD1 binding event phenotype enrichment analysis

To determine the significance of the overlap between TEAD1 binding events and GWAS-derived disease- and phenotype-associated genetic variants, we generated a custom GWAS catalogue for each phenotype in an ancestry specific manner. To this end, we downloaded the Genome Wide Association Studies Catalogue (https://www.ebi.ac.uk/gwas/) v1.0.2, as queried on June 20th 2019. Independent risk loci for each disease/phenotype were identified based on linkage disequilibrium (LD) pruning ($r2 < 0.2$). Risk loci across these independent genetic risk variants were identified by linkage disequilibrium expansion ($r2>0.8$) based on 1000 Genomes Data using PLINK (v.1.90b). This created a list of disease risk loci, along with the corresponding genetic variants within the LD block. This list of variants was then used for RELI analyses as previously published (Harley et al., 2018). Phenotypes were considered “TEAD1 loss specific” if: (1) they had 3 or more overlaps with TEAD1 binding event losses; (2) this overlap was significant according to RELI (corrected $p<0.01$); (3) this overlap was not significant for unchanged TEAD1 peaks ($p>0.01$), and (4) TEAD1 binding loss events were at least 2-fold enriched according to RELI. The three phenotypes meetings these criteria are presented in the Results, with further data included in **Supplemental Dataset 14.**

Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE# GSE254741 (reviewer login information: uhclkwceblejrep). A UCSC Genome Browser session for the hg19 genome build is available at http://genome.ucsc.edu/s/Ledsall/CMV_genomics.

Acknowledgments

This research was funded by National Institutes of Health (NIH) R01 HG010730, R01 GM055479, U01 AI130830 to M.T.W.; R01 AR073228, R01 NS099068, and R01 AI024717 to M.T.W. and L.C.K.; R01 AI148276, U01 HG011172, and U19 AI070235, and P30 AR070549 to L.C.K.; T32 ES007250 to A.V.H.; R21 DE026267, R01 AI121028 to W.E.M; F32 AI172329 and T32 AI007245 to L.A.M-N; R01 AI164709 to B.E.G; US Department of Veterans Affairs, I01 BX003850 to K.M.K; US Department of Veterans Affairs, I01 BX001834 and I01 BX006254 to J.B.H. We would like to thank Jaimie Miser for her assistance with culture work, Phillip Dexheimer for his help with the virus expression database (https://vexd.cchmc.org/), the Cincinnati Children's Hospital Medical Center Genomics Sequencing Facility (RRID SCR_022630) for high-throughput sequencing, the support of the Informatics Shared Facility in Information Services for Research (IS4R) at Cincinnati Children's Hospital Medical Center (RRID:SCR_022622), and Kevin Ernst for help with database management and computational support.

Author contributions

Project conception: M.T.W, L.C.K, J.B.H, and K.S; W.E.M provided TB40/E stocks, M.J.B, W.E.M, J.W., and K.S performed cell cultures and initiated viral infections; K.S conducted most experiments; M.H. and C.F. provided experimental support; S.P performed all initial informatics and QC analysis; K.S, S.P, L.A.M.-N., L.C.K, and M.T.W analyzed and interpreted data; S.P, A.V, L.E.E, O.D, and M.T.W designed and wrote informatics pipelines used for data analysis; M.M.L.L and B.Z performed the H3K27ac HiChIP experiment; L.E.E performed HiChIP analysis; O.A.D quantified and pooled the sequencing libraries for ATAC-seq, ChIP-seq and HiChIP-seq; A.C performed alternative splicing analysis; S.R assisted with qRT-PCRs; K.K provided critical insights at the early stages of the project and helped with data analysis; K.S, L.C.K, and M.T.W

wrote the first draft of the manuscript; K.S, S.P, L.E.E, L.C.K, and M.T.W finalized the manuscript with comments from all other authors.

Declaration of interests

All authors declare no conflicts of interests related to this manuscript.

Tables

Chr	Start	End	Number of TEAD1 loss events	Gene
chr15	67,200,000	67,500,000	19	*SMAD3*
chr14	51,800,000	52,100,000	17	*FRMD6*
chr3	30,500,000	30,800,000	14	*TGFBR2*
chr11	12,500,000	12,800,000	11	*TEAD1*
chr8	128,600,000	128,900,000	11	*MYC*
chr1	218,400,000	218,700,000	10	*TGFB2*
chr12	1,600,000	1,900,000	10	*WNT5B*
chr19	47,600,000	47,900,000	10	*BBC3*
chr5	141,900,000	142,200,000	10	*FGF1*
chr11	111,500,000	111,800,000	9	*PPP2R1B*
chr2	202,700,000	203,000,000	9	*FZD7*
chr6	7,600,000	7,900,000	9	*BMP6*
chr4	126,100,000	126,400,000	8	*FAT4*
chr7	100,600,000	100,900,000	8	*SERPINE1*
chr7	116,800,000	117,100,000	8	*WNT2*
chr8	28,200,000	28,500,000	8	*FZD3*
chr11	69,300,000	69,600,000	7	*CCND1*
chr17	42,500,000	42,800,000	7	*FZD2*
chr20	33,900,000	34,200,000	7	*GDF5*
chr3	55,400,000	55,700,000	7	*WNT5A*
chr6	132,100,000	132,400,000	7	*CTGF*

chr17	73,400,000	73,700,000	6	*LLGL2*
chr2	208,500,000	208,800,000	6	*FZD5*
chr20	49,200,000	49,500,000	6	*PARD6B*
chr7	5,400,000	5,700,000	6	*ACTB*
chr1	39,800,000	40,100,000	5	*BMP8A*
chr17	79,300,000	79,600,000	5	*ACTG1*
chr22	46,200,000	46,500,000	5	*WNT7B*
chr4	95,500,000	95,800,000	5	*BMPR1B*
chr5	167,600,000	167,900,000	5	*WWC1*

Table 1: Genomic regions with extensive loss of TEAD1 binding events upon HCMV infection. In an unbiased analysis, a 300kb window was drawn around the transcription start site of each gene with differential expression upon HCMV infection. The number of TEAD1 binding loss events within this window was then counted. All results with 5 or more TEAD1 loss events are provided in this table. Many of these genes encode members of the Hippo, TGF-beta, and WNT signaling pathways (see text).

Figure legends

Figure 1. Experimental design.

Schematic overview of the experimental design. **A.** Human foreskin fibroblasts or human retinal epithelial cells were infected with the TB40/E strain of Human Cytomegalovirus (HCMV) at a multiplicity of infection of 5 and 10, respectively. Uninfected and HCMV infected cells were harvested 48 hours post infection. **B.** Gene expression, chromatin accessibility, histone marks of active regulatory elements (H3K27ac), transcription factor occupancy (TEAD1 and CTCF), and chromatin looping were measured genome-wide using RNA-seq, ATAC-seq, ChIP-seq, and HiChIP, respectively. Differential analyses were employed to identify HCMV-dependent functional events on a genome-wide scale.

Figure 2. Extensive HCMV-mediated alterations to human chromatin accessibility.

A and **B.** Venn-diagram comparing ATAC-seq peaks in uninfected vs. HCMV infected fibroblasts (**A**) and retinal epithelial cells (**B**). **C** and **D.** ATAC-seq signal comparison in uninfected and HCMV infected fibroblasts (**C**) and retinal epithelial cells (**D**).

Figure 3. HCMV infection alters the accessibility of chromatin containing TEAD DNA binding motifs and avoids altering CTCF-containing sites.

Systematic prediction of human transcription factors with HCMV-altered binding. **A**. Left panel: transcription factor binding site motif enrichment comparison within ATAC-seq peaks that are unchanged with infection (X-axis) vs. peaks that are closed with infection (Y-axis). Right panel: same analysis comparing peaks that are unchanged with infection (X-axis) and peaks that are opened with infection (Y-axis). Each dot represents a human transcription factor binding site motif. Motifs are color-coded by TF family (see key). **B.** Same analysis in retinal epithelial cells. **C**. Percent of predicted binding sites for TEAD and CTCF in ATAC-seq peak regions unchanged with infection and regions closed by HCMV infection. **D**. Same analysis in retinal epithelial cells.

Figure 4. HCMV infection leads to widespread coincident loss of chromatin accessibility, TEAD1 binding, H3K27ac marks, and chromatin looping

A. HiChIP and ChIP-seq signal in the context of differentially accessible chromatin regions (ATAC-seq). Regions are split into those containing ATAC-seq signal that is unchanged (top), closed upon infection (middle), or opened upon infection (bottom). The corresponding normalized reads counts are depicted for (left to right): ATAC-seq, HiChIP, and ChIP-seq for TEAD1, CTCF, and H3K27ac. Each row in the heatmaps represents a single genomic locus. **B.** Genome browser images showing depletion of TEAD1, H3K27ac marks, and chromatin looping interactions proximal to Hippo pathway genes *FRMD6* and *RASSF2*. Solid boxes highlight differential TEAD1 binding sites. The *FRMD6* dashed box highlights a promoter and the *RASSF2* dashed box highlights an enhancer. Chromatin looping interactions lost upon infection are highlighted in blue.

Figure 5. HCMV infection alters Hippo signaling gene and protein expression levels

A. Differentially expressed genes in fibroblasts with HCMV infection. **B.** KEGG pathway enrichment analysis of differentially expressed genes. Pathways relevant to infection are highlighted with blue boxes. Key developmental pathways are highlighted with red boxes. **C.** Differentially expressed genes within the Hippo pathway. **D.** Gene expression profiles (transcripts per million [TPM] values) of all four TEAD family transcription factors with and without HCMV infection. **E.** Western blots of established TEAD1 targets THBS1 and CCN1 using whole cell lysates of uninfected and HCMV infected cells. GAPDH is used as control.

Figure 6. HCMV impairs TEAD1 activity through multiple distinct mechanisms

A. Representative Western blots for HCMV proteins IE1/2, YAP1, pYAP1, TEAD1 and the H3K27ac mark from whole cell lysates of uninfected and HCMV infected cells. GAPDH was used as a loading control. Additional Western blots (biological triplicates) are provided in **Supplemental Figure S4**, along with quantifications and p-values. **B.** Western blots using cytosolic and nuclear fractions obtained from uninfected and HCMV infected cells indicating the localization of YAP1 and pYAP1. GAPDH and Histone H3 were used as controls for cytoplasmic and nuclear fractions, respectively. **C.** Agarose gel image of RT-PCR products of TEAD1 exon-6 splicing events. The full length *TEAD1* targeted region is 91 bp long; it is 79 bp without exon-6. Jurkat cells, which have approximately equal expression of TEAD1 with and without exon 6 (Choi et al., 2022), were used as a control. **D.** Model depicting four distinct mechanisms by

which HCMV reduces the activity of the TEAD1 transcription factor. **E.** Enrichment of phenotype-associated genetic variants at HCMV-altered TEAD1 binding events. Enrichment values calculated by the RELI algorithm are presented for TEAD1 binding events for the lobe attachment (A) and ocular cup (B) phenotypes. Enrichment for TEAD1 binding loss with HCMV infection is statistically significant for each assessment except for the group for lobe attachment (shown in purple). For each bar, dots represent RELI enrichment results from different ancestral groups as defined in the original GWAS studies.

Supplemental Figure Legends

Supplemental Figure S1. Principle component analysis of ATAC-seq replicates of uninfected and HCMV infected fibroblasts and retinal epithelial cells.

Supplemental Figure S2. Heatmap of ChIP-seq peaks for TEAD1, CTCF, and H3K27ac between replicates in uninfected and HCMV infected fibroblasts.

Supplemental Figure S3. Heat map of gene expression profiles between replicates of uninfected and HCMV infected fibroblasts.

Supplemental Figure S4. Western blots of uninfected and HCMV infected fibroblasts in biological triplicates for TEAD1, H3K27ac, YAP1, pYAP1, and HCMV Immediate Early proteins IE1/2. For the HCMV IE1/2 western blot, β-Actin (red color) was used as a loading control. For TEAD1, H3K27ac, YAP1, and pYAP1, GAPDH was used as a control (red color). Quantification of Western blot signal intensities for each protein are shown next to the blot images. Signal quantification was performed using Emperia Studio software and p-values were calculated using two-tailed t-test on GraphPad Prism software. NS, not significant.

Supplemental Figure S5. A. Visualization of the two significant TEAD1 splicing changes in HCMV infected and uninfected cells. The splicing model of these events shows the percent spliced in (PSI) values, reported as percentages, for uninfected and HMCV infected cells. The red splice junction indicates the splicing outcome that is upregulated in HCMV infection. **B.** Sashimi plots depicting TEAD1 splice junctions and read coverage from exon 4 to exon 7, with the significant splicing changes of exon 6 (E6) skipping and partial inclusion of intron 5 (I5) highlighted in grey **C.** DNA sequence flanking TEAD1 exon 6 (green). RT-PCR primers are depicted in red.

Supplemental Figure S6. Comparison of HiChIP signal between replicates of uninfected and HCMV infected cells. Axes represent genomic bin location.

Supplemental Datasets

Supplemental Dataset 1. ATAC-seq data QC results.

Supplemental Dataset 2. ATAC-seq differential peak analysis results.

Supplemental Dataset 3. Chromatin looping interaction statistics.

Supplemental Dataset 4. HiChIP data QC results.

Supplemental Dataset 5. Pairwise statistical comparison of datasets generated in this study using RELI.

Supplemental Dataset 6. Chromatin looping event coordinates.

Supplemental Dataset 7. ChIP-seq data QC results.

Supplemental Dataset 8. ChIP-seq differential peak analysis results.

Supplemental Dataset 9. RNA-seq data QC results.

Supplemental Dataset 10. HCMV gene expression levels (TPMs).

Supplemental Dataset 11. RNA-seq DEG analysis results.

Supplemental Dataset 12. Regions of intense TEAD1 binding activity loss near differentially expressed genes.

Supplemental Dataset 13. Hippo pathway genes from KEGG.

Supplemental Dataset 14. Phenotype-associated genetic variant overlap analysis results (RELI algorithm applied to GWAS).

References

Adhikari, K., Reales, G., Smith, A.J., Konka, E., Palmen, J., Quinto-Sanchez, M., Acuna-Alonzo, V., Jaramillo, C., Arias, W., Fuentes, M.*, et al.* (2015). A genome-wide association study identifies multiple loci for variation in human ear morphology. Nat Commun *6*, 7500.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data.

Bhattacharyya, S., Chandra, V., Vijayanand, P., and Ay, F. (2019). Identification of significant chromatin contacts from HiChIP data by FitHiChIP. Nat Commun *10*, 4221.

Bonnemaijer, P.W.M., Leeuwen, E.M.V., Iglesias, A.I., Gharahkhani, P., Vitart, V., Khawaja, A.P., Simcoe, M., Hohn, R., Cree, A.J., Igo, R.P.*, et al.* (2019). Multi-trait genome-wide association study identifies new loci associated with optic disc parameters. Commun Biol *2*, 435.

Bradford, R.D., Yoo, Y.G., Golemac, M., Pugel, E.P., Jonjic, S., and Britt, W.J. (2015). Murine CMV-induced hearing loss is associated with inner ear inflammation and loss of spiral ganglia neurons. PLoS Pathog *11*, e1004774.

Campbell, M., Chantarasrivong, C., Yanagihashi, Y., Inagaki, T., Davis, R.R., Nakano, K., Kumar, A., Tepper, C.G., and Izumiya, Y. (2022). KSHV Topologically Associating Domains in Latent and Reactivated Viral Chromatin. J Virol *96*, e0056522.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Choi, J.A., Kim, J.E., Noh, S.J., Kyoung Kim, E., Park, C.K., and Paik, S.Y. (2017). Enhanced cytomegalovirus infection in human trabecular meshwork cells and its implication in glaucoma pathogenesis. Sci Rep *7*, 43349.

Choi, S., Lee, H.S., Cho, N., Kim, I., Cheon, S., Park, C., Kim, E.M., Kim, W., and Kim, K.K. (2022). RBFOX2-regulated TEAD1 alternative splicing plays a pivotal role in Hippo-YAP signaling. Nucleic Acids Res *50*, 8658-8673.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B.*, et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods *14*, 959-962.

Currey, L., Thor, S., and Piper, M. (2021). TEAD family transcription factors in development and disease. Development *148*.

Daley, T., and Smith, A.D. (2013). Predicting the molecular complexity of sequencing libraries. Nat Methods *10*, 325-327.

Deng, Y., Wu, A., Li, P., Li, G., Qin, L., Song, H., and Mak, K.K. (2016). Yap1 Regulates Multiple Steps of Chondrocyte Differentiation during Skeletal Development and Bone Repair. Cell Rep *14*, 2224-2237.

Di Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. Nature Biotechnology *35*, 316-319.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst *3*, 95-98.

Emig, D., Salomonis, N., Baumbach, J., Lengauer, T., Conklin, B.R., and Albrecht, M. (2010). AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. Nucleic Acids Res *38*, W755-762.

Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P., and Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. Nature Biotechnology *38*, 276-278.

Fontan, C.T., James, C.D., Prabhakar, A.T., Bristol, M.L., Otoa, R., Wang, X., Karimi, E., Rajagopalan, P., Basu, D., and Morgan, I.M. (2022). A Critical Role for p53 during the HPV16 Life Cycle. Microbiol Spectr *10*, e0068122.

Forte, E., Ayaloglu Butun, F., Marinaccio, C., Schipma, M.J., Piunti, A., Schroeder, M.W., Kandpal, M., Shilatifard, A., Abecassis, M., and Hummel, M. (2021). Epigenetic reprogramming of host and viral genes by Human Cytomegalovirus infection in Kasumi-3 myeloid progenitor cells at early times post-infection. J Virol.

Forte, E., Zhang, Z., Thorp, E.B., and Hummel, M. (2020). Cytomegalovirus Latency and Reactivation: An Intricate Interplay With the Host Immune Response. Front Cell Infect Microbiol *10*, 130.

Fortunato, E.A. (2021). Using Diploid Human Fibroblasts as a Model System to Culture, Grow, and Study Human Cytomegalovirus Infection. Methods Mol Biol *2244*, 39-50.

Fowler, K.B., and Boppana, S.B. (2006). Congenital cytomegalovirus (CMV) infection and hearing deficit. J Clin Virol *35*, 226-231.

Frum, T., Murphy, T.M., and Ralston, A. (2018). HIPPO signaling resolves embryonic cell fate conflicts during establishment of pluripotency in vivo. Elife *7*.

Gharahkhani, P., Burdon, K.P., Cooke Bailey, J.N., Hewitt, A.W., Law, M.H., Pasquale, L.R., Kang, J.H., Haines, J.L., Souzeau, E., Zhou, T.*, et al.* (2018). Analysis combining correlated glaucoma traits identifies five new risk loci for open-angle glaucoma. Sci Rep *8*, 3124.

Gnedeva, K., Wang, X., McGovern, M.M., Barton, M., Tao, L., Trecek, T., Monroe, T.O., Llamas, J., Makmura, W., Martin, J.F.*, et al.* (2020). Organ of Corti size is governed by Yap/Tead-mediated progenitor self-renewal. Proc Natl Acad Sci U S A *117*, 13552-13561.

Goderis, J., De Leenheer, E., Smets, K., Van Hoecke, H., Keymeulen, A., and Dhooge, I. (2014). Hearing loss and congenital CMV infection: a systematic review. Pediatrics *134*, 972-982.

Harley, J.B., Chen, X., Pujato, M., Miller, D., Maddox, A., Forney, C., Magnusen, A.F., Lynch, A., Chetal, K., Yukawa, M.*, et al.* (2018). Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. Nat Genet *50*, 699-707.

Harshil, P., Ewels, P., Peltzer, A., Hammarén, R., Botvinnik, O., Sturm, G., Moreno, D., Pranathi, V., Silviamorins, Pantano, L.*, et al.* (2022). nf-core/rnaseq: nf-core/rnaseq v3.8.1 - Plastered Magnesium Mongoose (Zenodo).

He, C., Mao, D., Hua, G., Lv, X., Chen, X., Angeletti, P.C., Dong, J., Remmenga, S.W., Rodabaugh, K.J., Zhou, J.*, et al.* (2015). The Hippo/YAP pathway interacts with EGFR signaling and HPV oncoproteins to regulate cervical cancer progression. EMBO Mol Med *7*, 1426-1449.

Hein, M.Y., and Weissman, J.S. (2021). Functional single-cell genomics of human cytomegalovirus infection. Nat Biotechnol.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Hertel, L., and Mocarski, E.S. (2004). Global analysis of host cell gene expression late during cytomegalovirus infection reveals extensive dysregulation of cell cycle gene expression and induction of Pseudomitosis independent of US28 function. J Virol *78*, 11988-12011.

Hitz, B.C., Lee, J.W., Jolanki, O., Kagda, M.S., Graham, K., Sud, P., Gabdank, I., Seth Strattan, J., Sloan, C.A., Dreszer, T.*, et al.* (2023). The ENCODE Uniform Analysis Pipelines. bioRxiv.

Holden, J.K., and Cunningham, C.N. (2018). Targeting the Hippo Pathway and Cancer through the TEAD Family of Transcription Factors. Cancers (Basel) *10*.

Ijezie, E.C., O'Dowd, J.M., Kuan, M.I., Faeth, A.R., and Fortunato, E.A. (2023). HCMV Infection Reduces Nidogen-1 Expression, Contributing to Impaired Neural Rosette Development in Brain Organoids. J Virol *97*, e0171822.

Jiang, S., Zhou, H., Liang, J., Gerdt, C., Wang, C., Ke, L., Schmidt, S.C.S., Narita, Y., Ma, Y., Wang, S.*, et al.* (2017). The Epstein-Barr Virus Regulome in Lymphoblastoid Cells. Cell Host Microbe *22*, 561-573 e564.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006.

Kong, Q., Li, J., Zhao, L., Shi, P., Liu, X., Bian, C., Liu, J., and Liu, T. (2021). Human cytomegalovirus inhibits the proliferation and invasion of extravillous cytotrophoblasts via Hippo-YAP pathway. Virol J *18*, 214.

Kopylova, E., Noe, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics *28*, 3211-3217.

Krueger, F., James, F., Ewels, P., Afyounian, E., and Schuster-Boeckler, B. (2021). FelixKrueger/TrimGalore: v0.6.7 (Zenodo).

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A.*, et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res *44*, W90-97.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650-665.

Lee, C.H., and Grey, F. (2020). Systems Virology and Human Cytomegalovirus: Using High Throughput Approaches to Identify Novel Host-Virus Interactions During Lytic Infection. Front Cell Infect Microbiol *10*, 280.

Lee, J.W., Ottojolanki, Kim, D., Strattan, J.S., Anshul, K., Mawla, A.M., Nordström, K., and Annashcherbina (2021a). ENCODE-DCC/atac-seq-pipeline: v2.0.0 (Zenodo).

Lee, J.W., Strattan, J.S., Annashcherbina, Sebby, K., Kagda, M., and Maurizio, P.L. (2021b). ENCODE-DCC/chip-seq-pipeline2: v2.0.0 (Zenodo).

Lee, W., Mitchell, P., and Tjian, R. (1987). Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell *49*, 741-752.

Li, G., and Kamil, J.P. (2016). Viral Regulation of Cell Tropism in Human Cytomegalovirus. J Virol *90*, 626-629.

Lindstrom, S., Thompson, D.J., Paterson, A.D., Li, J., Gierach, G.L., Scott, C., Stone, J., Douglas, J.A., dos-Santos-Silva, I., Fernandez-Navarro, P.*, et al.* (2014). Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. Nat Commun *5*, 5303.

Lindstrom, S., Vachon, C.M., Li, J., Varghese, J., Thompson, D., Warren, R., Brown, J., Leyland, J., Audley, T., Wareham, N.J.*, et al.* (2011). Common variants in ZNF365 are associated with both mammographic density and breast cancer risk. Nat Genet *43*, 185-187.

Liu, G., Yu, F.X., Kim, Y.C., Meng, Z., Naipauer, J., Looney, D.J., Liu, X., Gutkind, J.S., Mesri, E.A., and Guan, K.L. (2015). Kaposi sarcoma-associated herpesvirus promotes tumorigenesis by modulating the Hippo pathway. Oncogene *34*, 3536-3546.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Luo, Y., Hitz, B.C., Gabdank, I., Hilton, J.A., Kagda, M.S., Lam, B., Myers, Z., Sud, P., Jou, J., Lin, K.*, et al.* (2020). New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. Nucleic Acids Res *48*, D882-D889.

Ma, S., Meng, Z., Chen, R., and Guan, K.L. (2019). The Hippo Pathway: Biology and Pathophysiology. Annu Rev Biochem *88*, 577-604.

Marti-Carreras, J., and Maes, P. (2019). Human cytomegalovirus genomics and transcriptomics through the lens of next-generation sequencing: revision and future challenges. Virus Genes *55*, 138-164.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j *17*, 10.

McKinney, C., Zavadil, J., Bianco, C., Shiflett, L., Brown, S., and Mohr, I. (2014). Global reprogramming of the cellular translational landscape facilitates cytomegalovirus replication. Cell Rep *6*, 9-17.

Meenakshi S. Kagda, B.L., Casey Litton, Corinn Small, Cricket A. Sloan,, Emma Spragins, F.T., Ian Whaling, Idan Gabdank, Ingrid Youngworth,, J. Seth Strattan, J.H., Jennifer Jou, Jessica Au, Jin-Wook Lee, Kalina, Andreeva, K.G., Khine Lin, Matt Simison, Otto Jolanki, Paul Sud,, Pedro Assis, P.A., Stuart Miyasato, Weiwei Zhong, Yunhai Luo,, and Zachary Myers, J.M.C.a.B.C.H. (2023). Data navigation on the ENCODE portal (arxiv.org).

Moon, K.H., and Kim, J.W. (2018). Hippo Signaling Circuit and Divergent Tissue Growth in Mammalian Eye. Mol Cells *41*, 257-263.

Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods *13*, 919-922.

Nichols, D.B., De Martini, W., and Cottrell, J. (2017). Poxviruses Utilize Multiple Strategies to Inhibit Apoptosis. Viruses *9*.

Nightingale, K., Lin, K.M., Ravenhill, B.J., Davies, C., Nobre, L., Fielding, C.A., Ruckova, E., Fletcher-Etherington, A., Soday, L., Nichols, H.*, et al.* (2018). High-Definition Analysis of Host Protein Stability during Human Cytomegalovirus Infection Reveals Antiviral Factors and Viral Evasion Mechanisms. Cell Host Microbe *24*, 447-460 e411.

Nogalski, M.T., Solovyov, A., Kulkarni, A.S., Desai, N., Oberstein, A., Levine, A.J., Ting, D.T., Shenk, T., and Greenbaum, B.D. (2019). A tumor-specific endogenous repetitive element is induced by herpesviruses. Nat Commun *10*, 90.

Nystad, M., Fagerheim, T., Brox, V., Fortunato, E.A., and Nilssen, O. (2008). Human cytomegalovirus (HCMV) and hearing impairment: infection of fibroblast cells with HCMV induces chromosome breaks at 1q23.3, between loci DFNA7 and DFNA49 -- both involved in dominantly inherited, sensorineural, hearing impairment. Mutat Res *637*, 56-65.

Oberstein, A., and Shenk, T. (2017). Cellular responses to human cytomegalovirus infection: Induction of a mesenchymal-to-epithelial transition (MET) phenotype. Proc Natl Acad Sci U S A *114*, E8244-E8253.

Okonechnikov, K., Conesa, A., and Garcia-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics *32*, 292-294.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods *14*, 417-419.

Pocaterra, A., Romani, P., and Dupont, S. (2020). YAP/TAZ functions and their regulation at a glance. J Cell Sci *133*.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Rafailidis, P.I., Mourtzoukou, E.G., Varbobitis, I.C., and Falagas, M.E. (2008). Severe cytomegalovirus infection in apparently immunocompetent patients: a systematic review. Virol J *5*, 47.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.

Rausch, V., and Hansen, C.G. (2020). The Hippo Pathway, YAP/TAZ, and the Plasma Membrane. Trends Cell Biol *30*, 32-48.

Rory Stark, G.B. (2017). DiffBind (Bioconductor).

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R.*, et al.* (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389-393.

Rothenburg, S., and Brennan, G. (2020). Species-Specific Host-Virus Interactions: Implications for Viral Host Range and Virulence. Trends Microbiol *28*, 46-56.

Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. Nat Rev Genet *19*, 789-800.

Ruckert, T., Lareau, C.A., Mashreghi, M.F., Ludwig, L.S., and Romagnani, C. (2022). Clonal expansion and epigenetic inheritance of long-lasting NK cell memory. Nat Immunol *23*, 1551-1563.

Sayols, S., Scherzinger, D., and Klein, H. (2016). dupRadar: a Bioconductor package for the assessment of PCR artifacts in RNA-Seq data. BMC Bioinformatics *17*, 428.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol *16*, 259.

Shaffer, J.R., Li, J., Lee, M.K., Roosenboom, J., Orlova, E., Adhikari, K., andMe Research, T., Gallo, C., Poletti, G., Schuler-Faccini, L.*, et al.* (2017). Multiethnic GWAS Reveals Polygenic Architecture of Earlobe Attachment. Am J Hum Genet *101*, 913-924.

Shi, C., Rattray, M., and Orozco, G. (2020). HiChIP-Peaks: a HiChIP peak calling algorithm. Bioinformatics *36*, 3625-3631.

Sieh, W., Rothstein, J.H., Klein, R.J., Alexeeff, S.E., Sakoda, L.C., Jorgenson, E., McBride, R.B., Graff, R.E., McGuire, V., Achacoso, N.*, et al.* (2020). Identification of 31 loci for mammographic density phenotypes and their associations with breast cancer risk. Nat Commun *11*, 5116.

Singh, D.R., Nelson, S.E., Pawelski, A.S., Kansra, A.S., Fogarty, S.A., Bristol, J.A., Ohashi, M., Johannsen, E.C., and Kenney, S.C. (2023a). Epstein-Barr virus LMP1 protein promotes proliferation and inhibits differentiation of epithelial cells via activation of YAP and TAZ. Proc Natl Acad Sci U S A *120*, e2219755120.

Singh, S., Maheshwari, A., and Boppana, S. (2023b). CMV-induced Hearing Loss. Newborn (Clarksville) *2*, 249-262.

Sinzger, C., Digel, M., and Jahn, G. (2008). Cytomegalovirus cell tropism. Curr Top Microbiol Immunol *325*, 63-83.

Sinzger, C., Grefte, A., Plachter, B., Gouw, A.S., The, T.H., and Jahn, G. (1995). Fibroblasts, epithelial cells, endothelial cells and smooth muscle cells are major targets of human cytomegalovirus infection in lung and gastrointestinal tissues. J Gen Virol *76 (Pt 4)*, 741-750.

Smeriglio, P., Wangsawihardja, F.V., Leu, R., and Mustapha, M. (2019). TSP1 and TSP2 Have Unique and Overlapping Roles in Protecting against Noise-Induced Auditory Synaptopathy. Neuroscience *408*, 68-80.

Springelkamp, H., Iglesias, A.I., Mishra, A., Hohn, R., Wojciechowski, R., Khawaja, A.P., Nag, A., Wang, Y.X., Wang, J.J., Cuellar-Partida, G.*, et al.* (2017). New insights into the genetics of primary open-angle glaucoma based on meta-analyses of intraocular pressure and optic disc characteristics. Hum Mol Genet *26*, 438-453.

Springelkamp, H., Mishra, A., Hysi, P.G., Gharahkhani, P., Hohn, R., Khor, C.C., Cooke Bailey, J.N., Luo, X., Ramdas, W.D., Vithana, E.*, et al.* (2015). Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. Genet Epidemiol *39*, 207-216.

Stein, C., Bardet, A.F., Roma, G., Bergling, S., Clay, I., Ruchti, A., Agarinis, C., Schmelzle, T., Bouwmeester, T., Schubeler, D.*, et al.* (2015). YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation of Enhancers. PLoS Genet *11*, e1005465.

Swanson, E.C., and Schleiss, M.R. (2013). Congenital cytomegalovirus infection: new prospects for prevention and therapy. Pediatr Clin North Am *60*, 335-349.

Team, R.C. (2023). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Totaro, A., Panciera, T., and Piccolo, S. (2018). YAP/TAZ upstream signals and downstream responses. Nat Cell Biol *20*, 888-899.

Van Damme, E., Thys, K., Tuefferd, M., Van Hove, C., Aerssens, J., and Van Loock, M. (2016). HCMV Displays a Unique Transcriptome of Immunomodulatory Genes in Primary Monocyte-Derived Cell Types. PLoS One *11*, e0164843.

van Den Pol, A.N., Mocarski, E., Saederup, N., Vieira, J., and Meier, T.J. (1999). Cytomegalovirus cell tropism, replication, and gene transfer in brain. J Neurosci *19*, 10948-10965.

Van Sciver, N., Ohashi, M., Pauly, N.P., Bristol, J.A., Nelson, S.E., Johannsen, E.C., and Kenney, S.C. (2021). Hippo signaling effectors YAP and TAZ induce Epstein-Barr Virus (EBV) lytic reactivation through TEADs in epithelial cells. PLoS Pathog *17*, e1009783.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184-2185.

Wang, M., Dong, Y., Gao, S., Zhong, Z., Cheng, C., Qiang, R., Zhang, Y., Shi, X., Qian, X., Gao, X.*, et al.* (2022). Hippo/YAP signaling pathway protects against neomycin-induced hair cell damage in the mouse cochlea. Cell Mol Life Sci *79*, 79.

Wang, Z., Lu, W., Zhang, Y., Zou, F., Jin, Z., and Zhao, T. (2019). The Hippo Pathway and Viral Infections. Front Microbiol *10*, 3033.

Weekes, M.P., Tomasec, P., Huttlin, E.L., Fielding, C.A., Nusinow, D., Stanton, R.J., Wang, E.C.Y., Aicheler, R., Murrell, I., Wilkinson, G.W.G.*, et al.* (2014). Quantitative temporal viromics: an approach to investigate host-pathogen interaction. Cell *157*, 1460-1472.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K.*, et al.* (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431-1443.

Wolf, B.K., Zhao, Y., McCray, A., Hawk, W.H., Deary, L.T., Sugiarto, N.W., LaCroix, I.S., Gerber, S.A., Cheng, C., and Wang, X. (2023). Cooperation of chromatin remodeling SWI/SNF complex and pioneer factor AP-1 shapes 3D enhancer landscapes. Nat Struct Mol Biol *30*, 10-21.

Xu, J., Liu, X., Zhang, X., Marshall, B., Dong, Z., Liu, Y., Espinosa-Heidmann, D.G., and Zhang, M. (2020). Ocular cytomegalovirus latency exacerbates the development of choroidal neovascularization. J Pathol *251*, 200-212.

Yamaguchi, N. (2020). Multiple Roles of Vestigial-Like Family Members in Tumor Development. Front Oncol *10*, 1266.

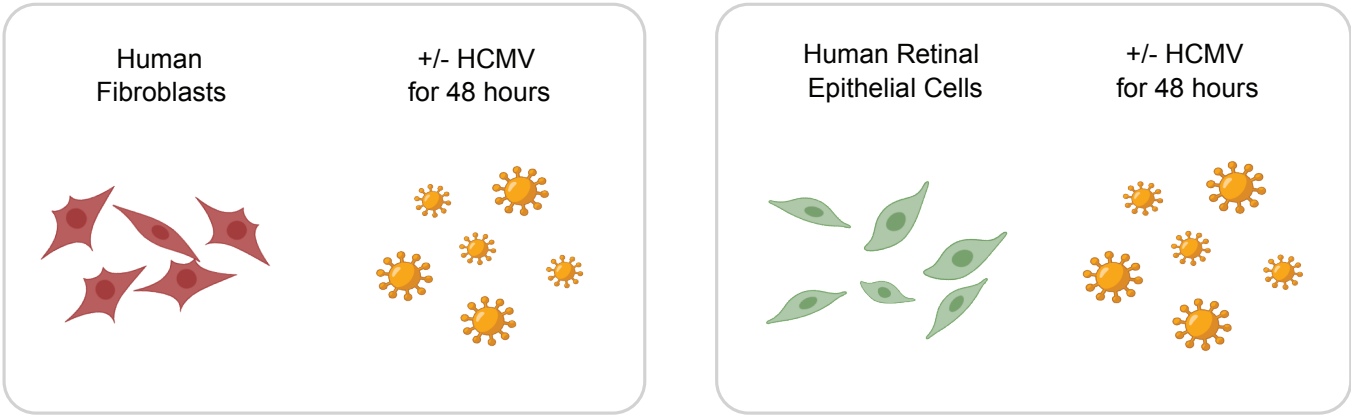
Zalckvar, E., Paulus, C., Tillo, D., Asbach-Nitzsche, A., Lubling, Y., Winterling, C., Strieder, N., Mucke, K., Goodrum, F., Segal, E.*, et al.* (2013). Nucleosome maps of the human cytomegalovirus genome reveal a temporal switch in chromatin organization linked to a major IE protein. Proc Natl Acad Sci U S A *110*, 13126-13131.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zuhair, M., Smit, G.S.A., Wallis, G., Jabbar, F., Smith, C., Devleesschauwer, B., and Griffiths, P. (2019). Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. Rev Med Virol *29*, e2034.

Figure 1

A. Human Cytomegalovirus (HCMV) infection system



B. Functional genomics of the human genome response to HCMV infection

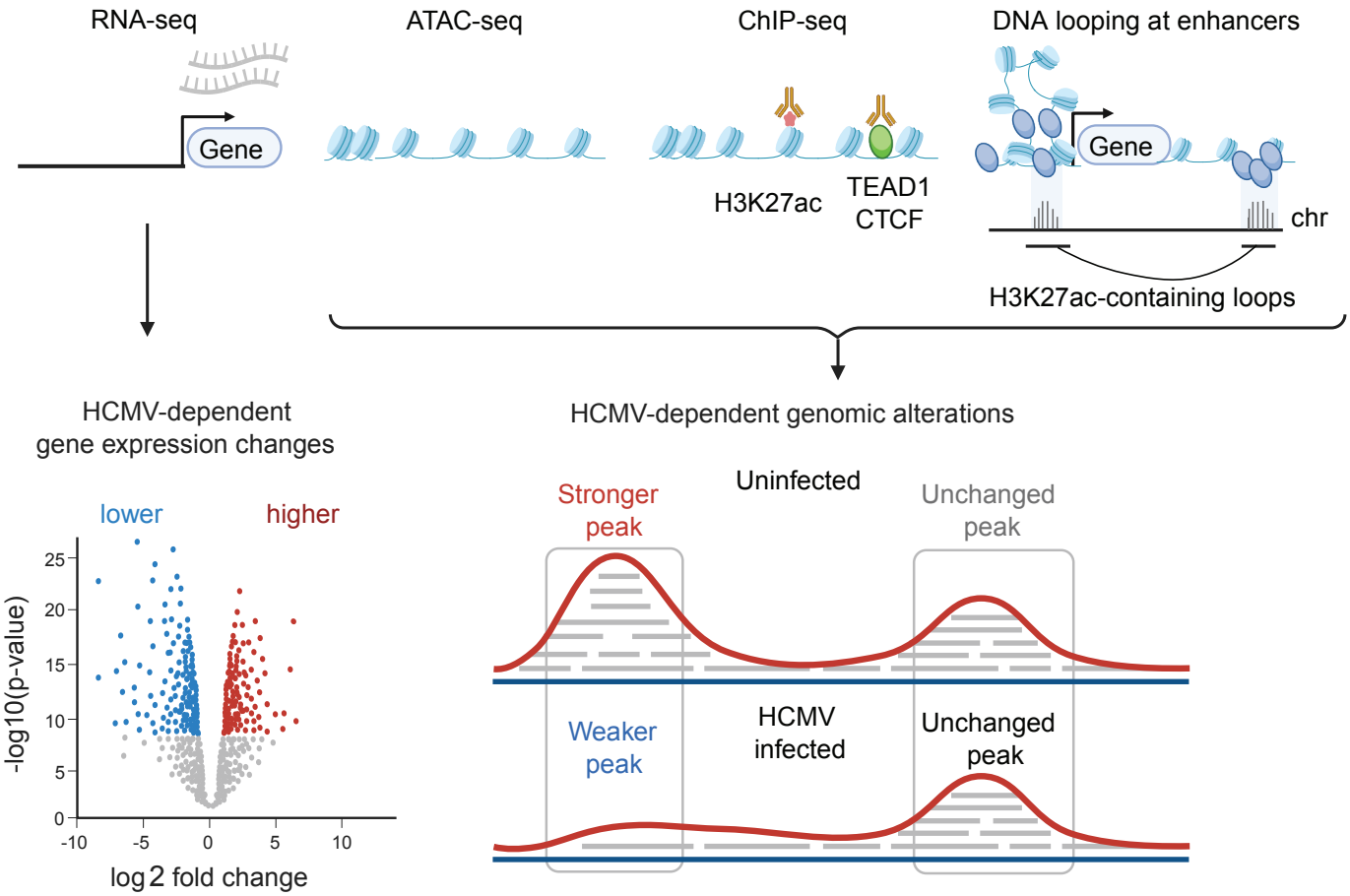
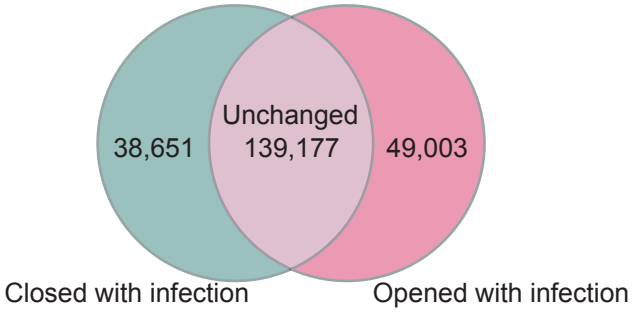
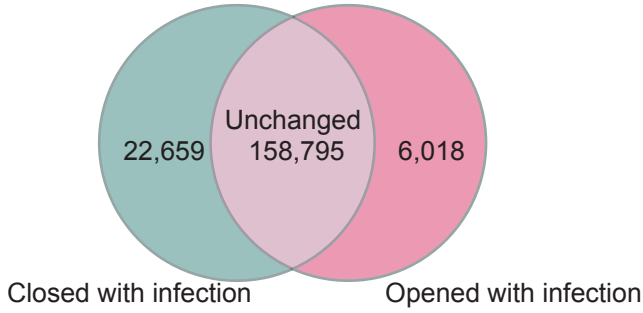


Figure 2

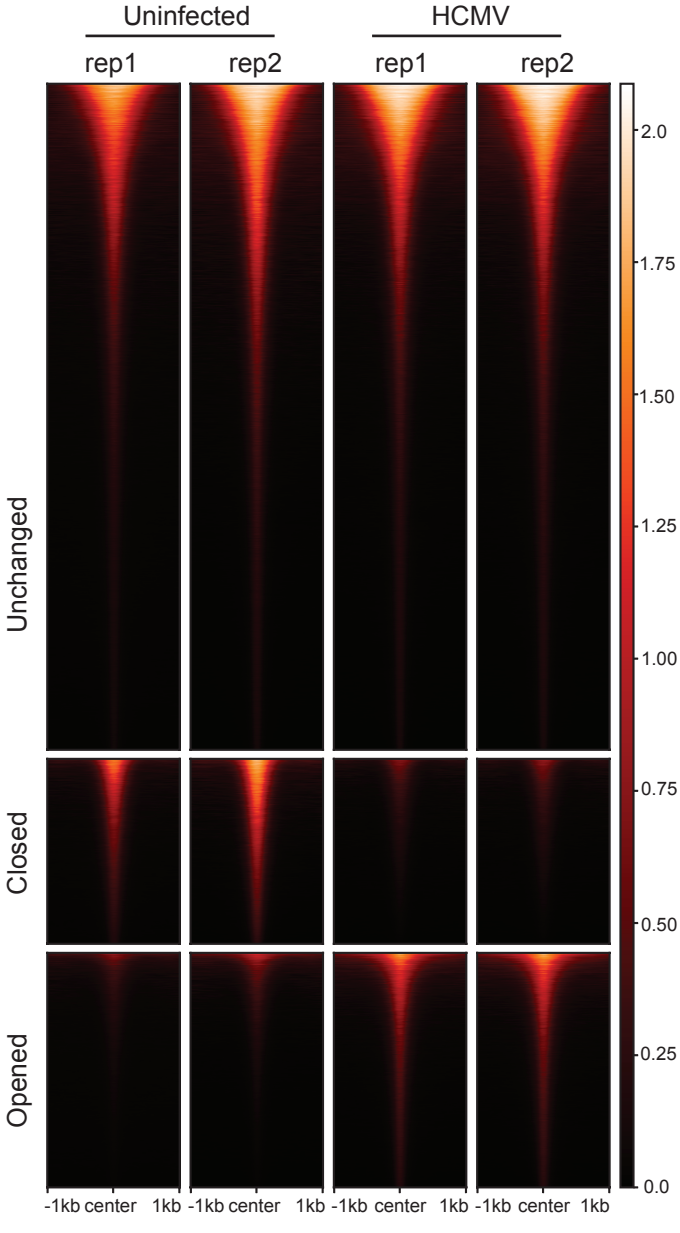
A. Chromatin accessibility - fibroblasts



B. Chromatin accessibility - retinal epithelial cells



C. ATAC-seq signal - fibroblasts



D. ATAC-seq signal - retinal epithelial cells

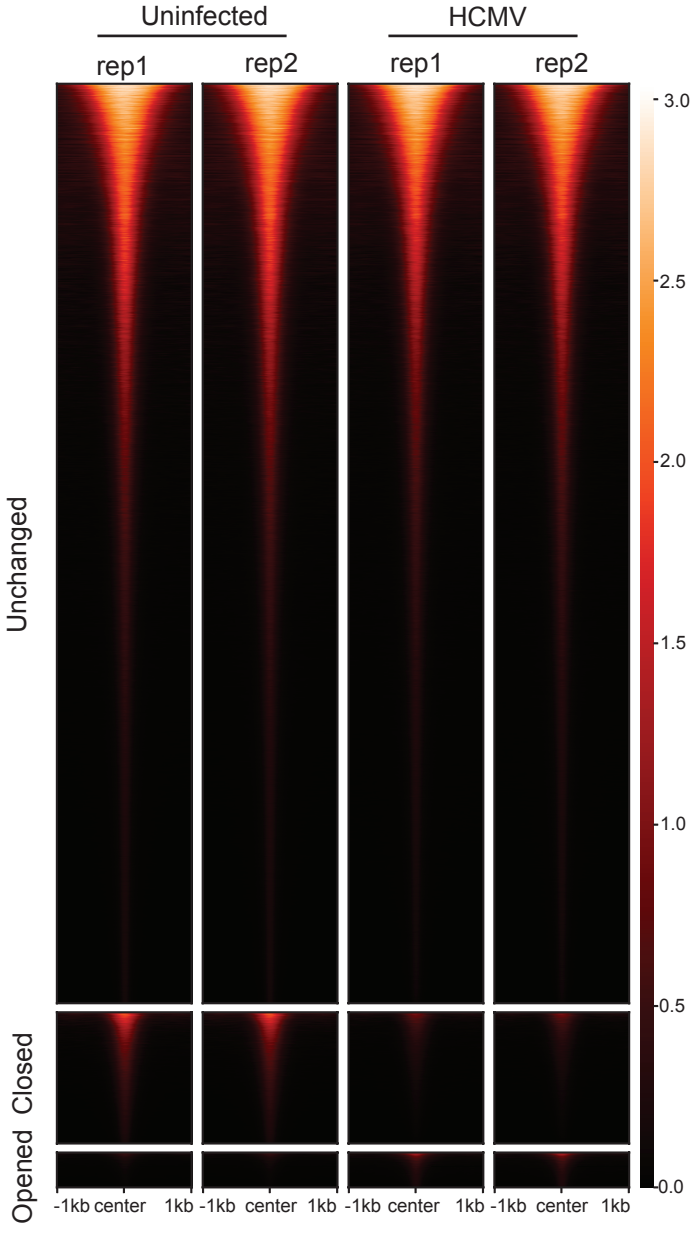
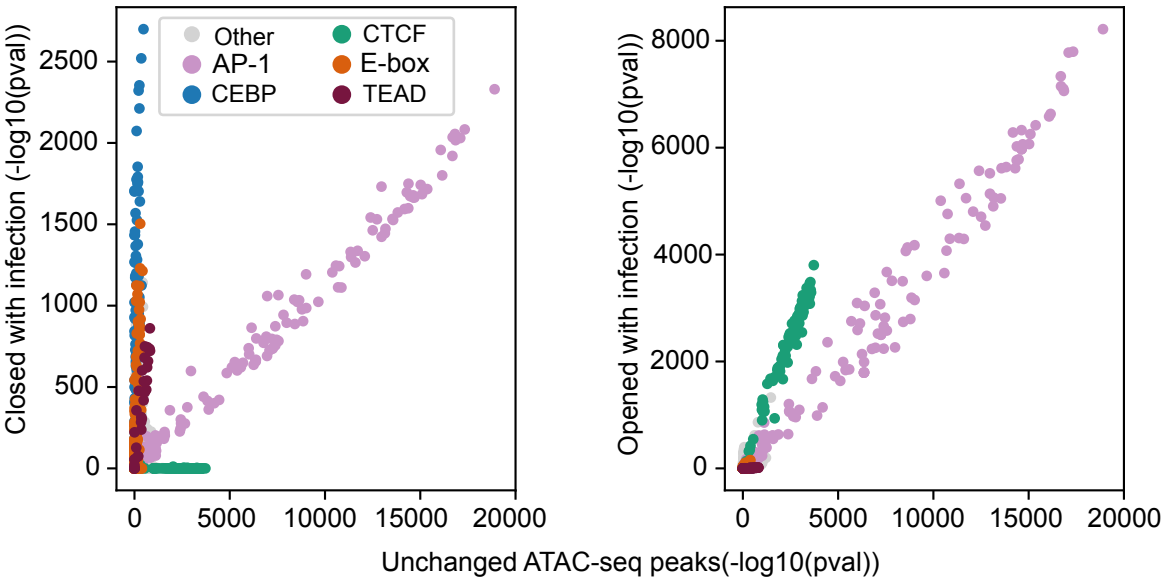
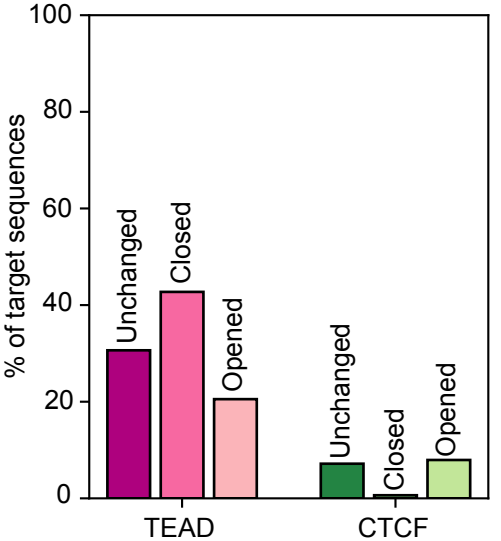


Figure 3

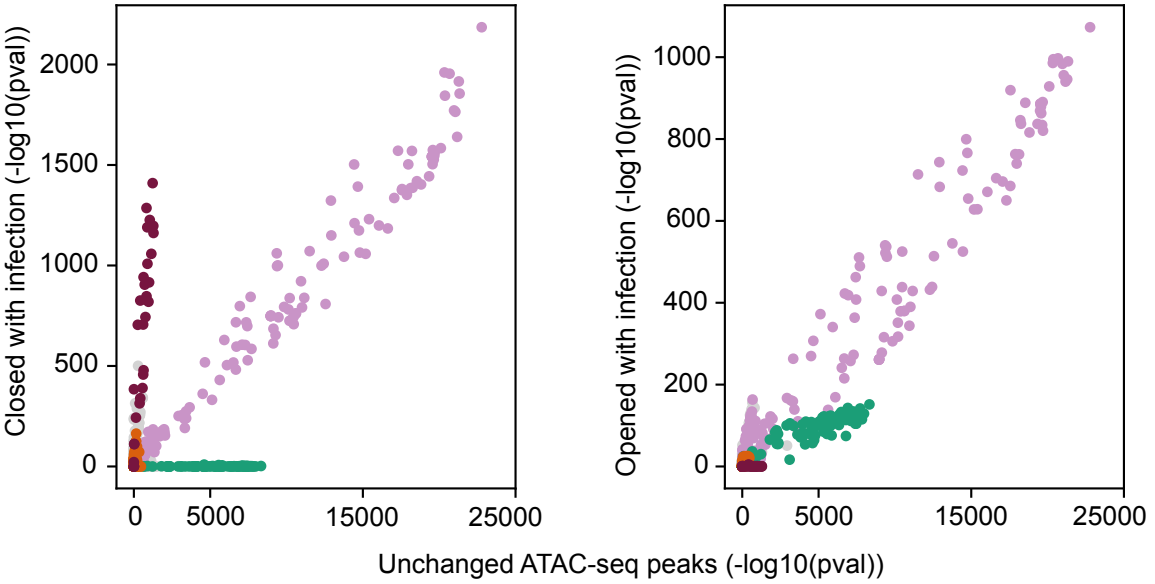
A. TF motif enrichment in fibroblasts



C. TF motif presence in fibroblasts



B. TF motif enrichment in retinal epithelial cells



D. TF motif presence in retinal epithelial cells

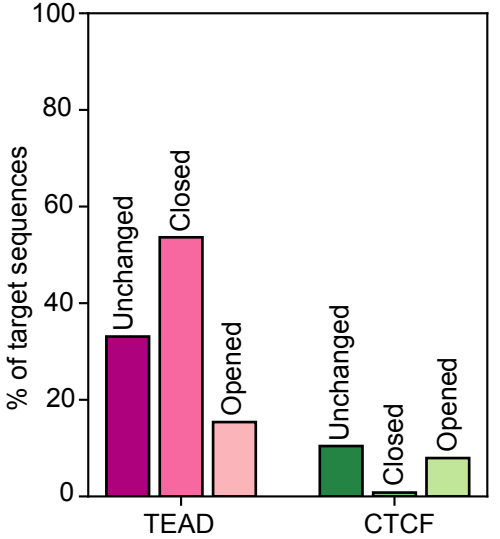
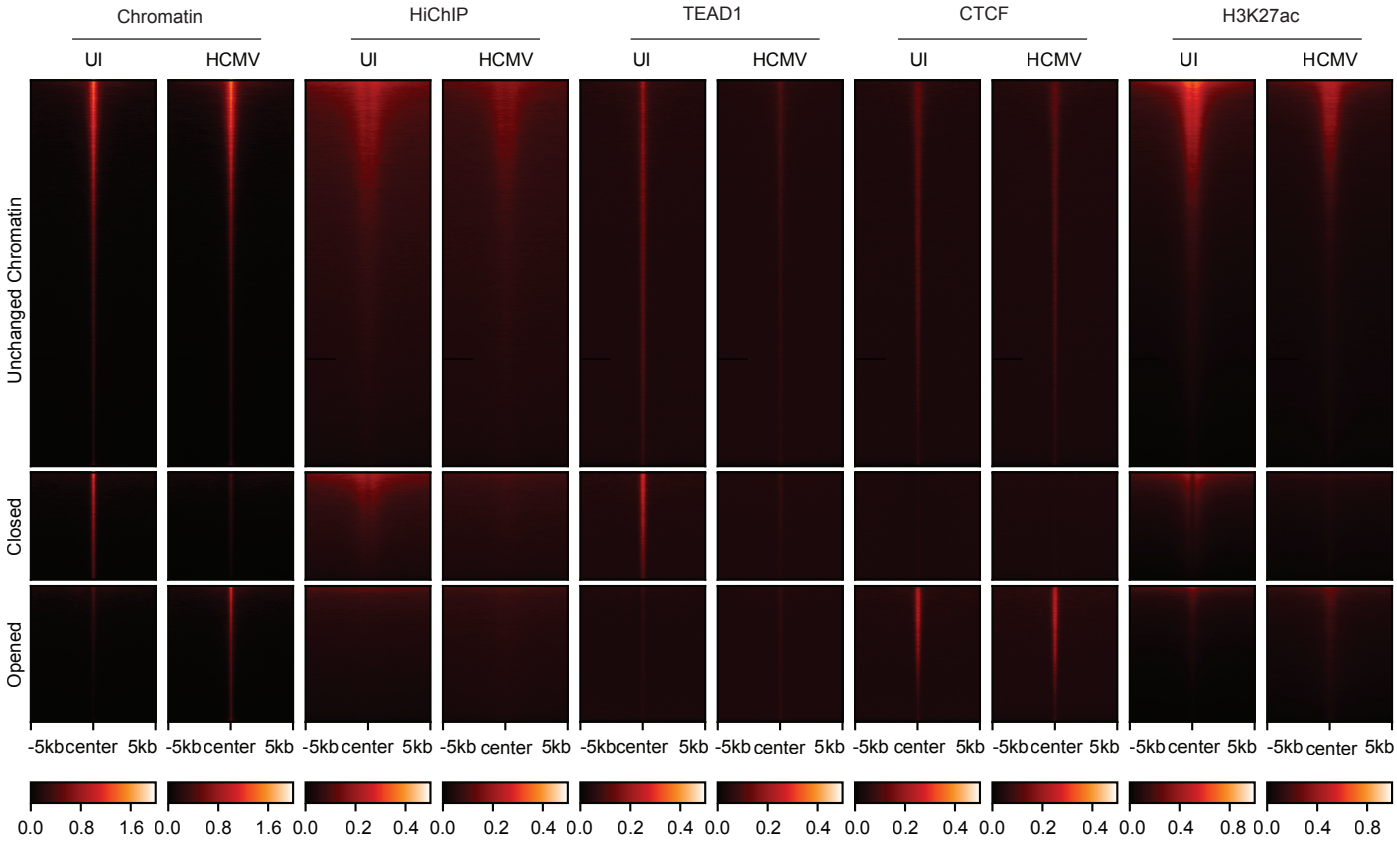


Figure 4

A. Normalized read depth at regions of HCMV-altered and unaltered chromatin on fibroblasts



B. Hippo pathway genes *FRMD6* and *RASSF2*

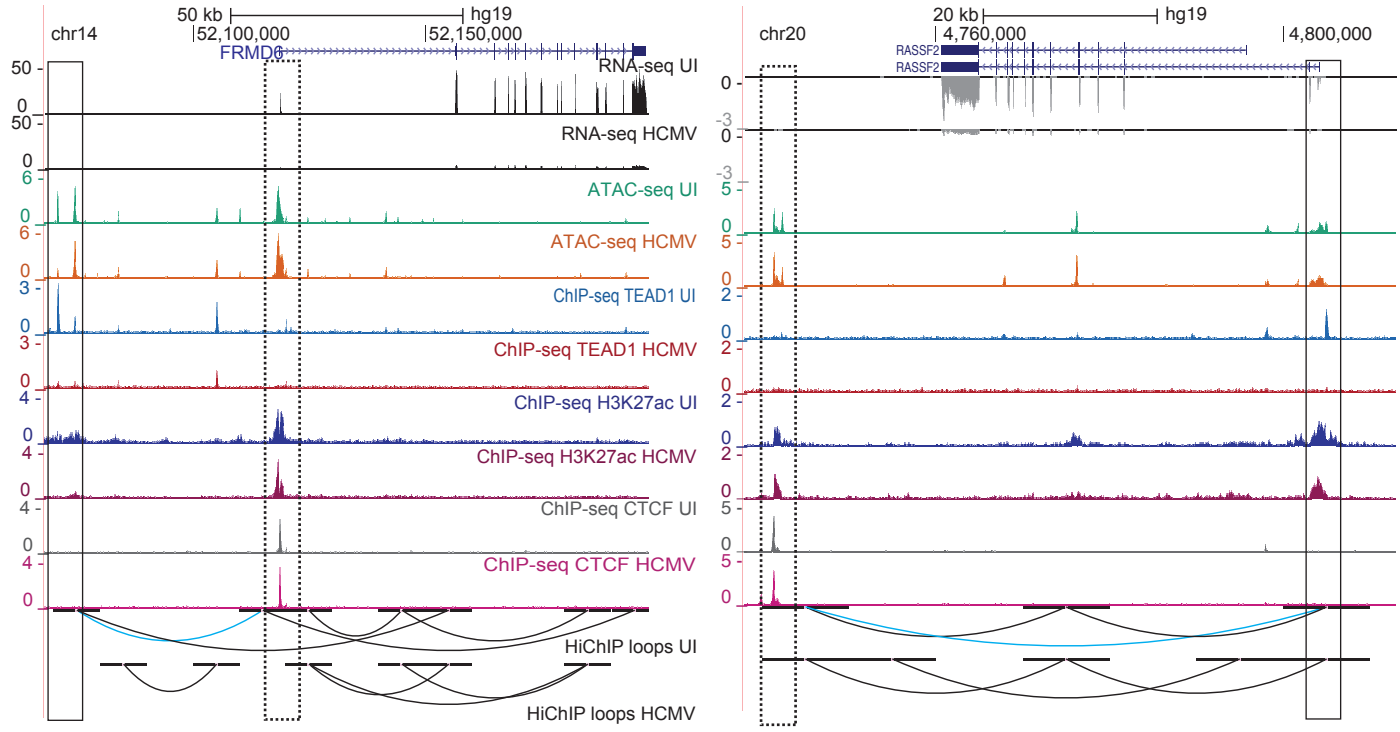
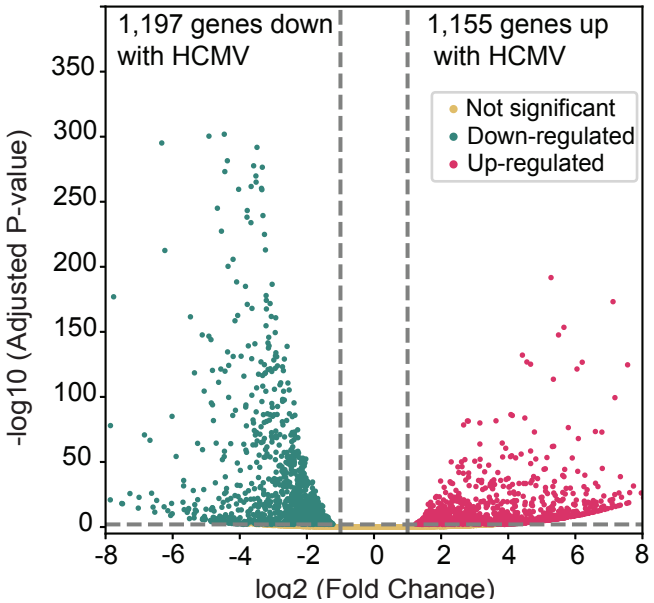
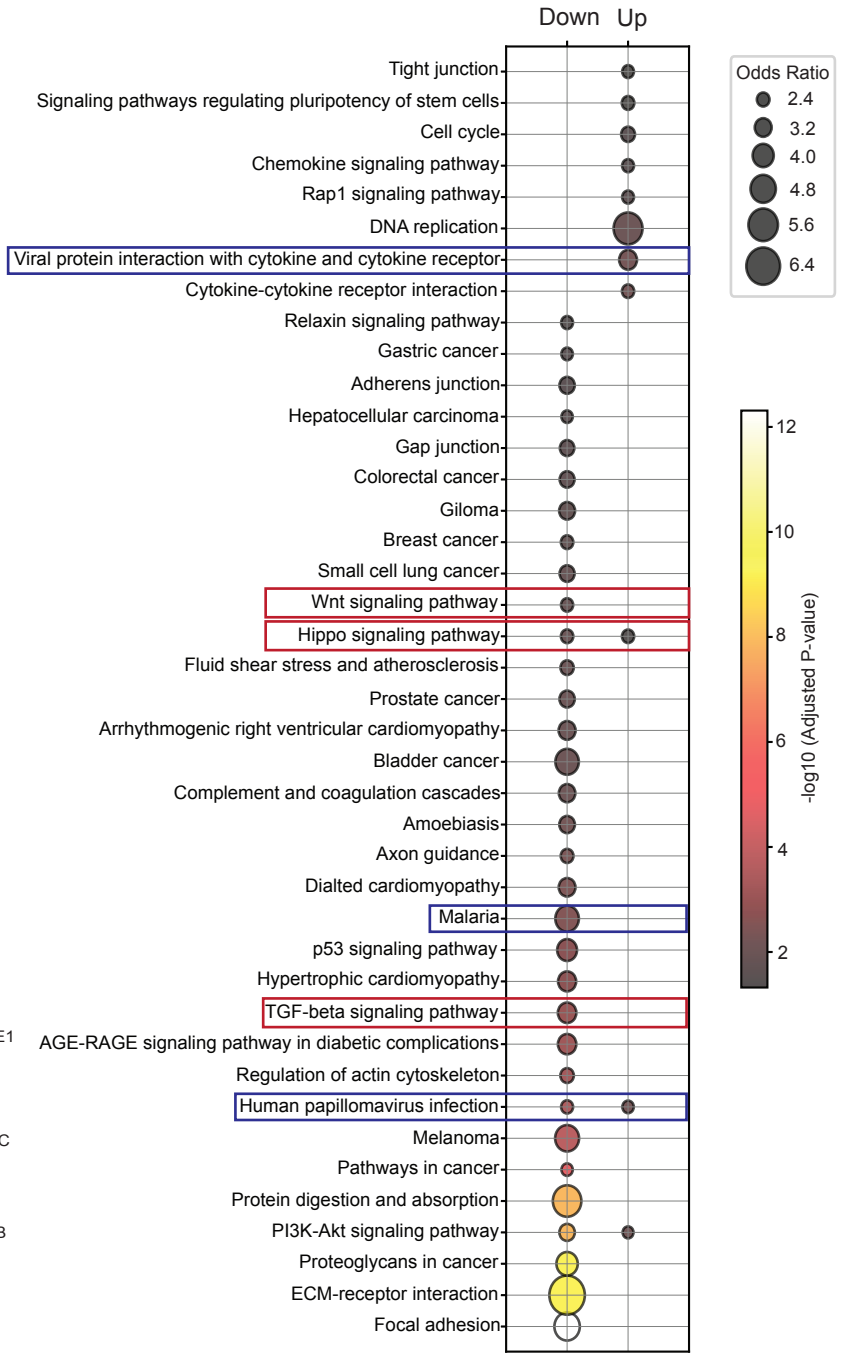


Figure 5

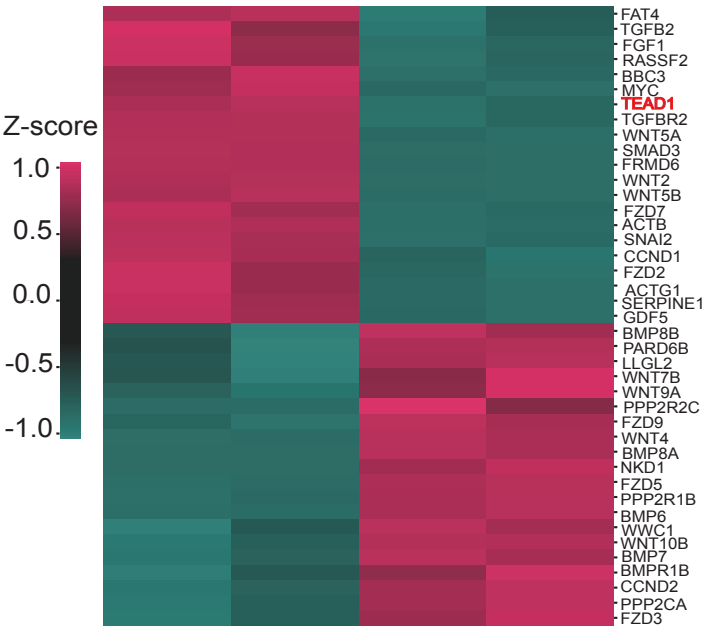
A. HCMV-dependent human gene expression in fibroblasts



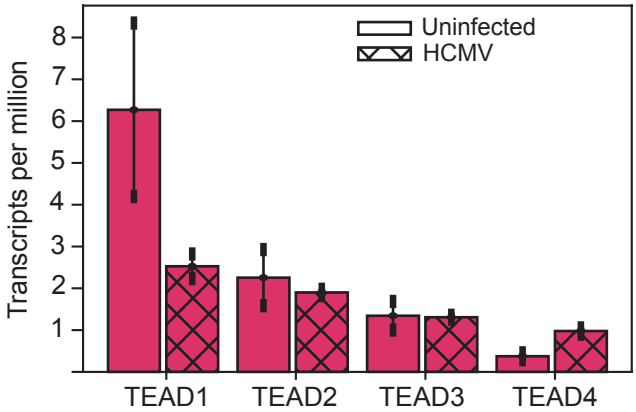
B. Pathway enrichment of HCMV-dependent human genes



C. Differentially expressed Hippo pathway genes



D. TEAD family gene expression levels



E. TEAD1 target gene protein expression levels

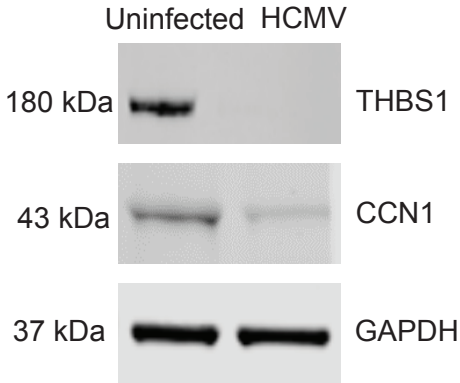
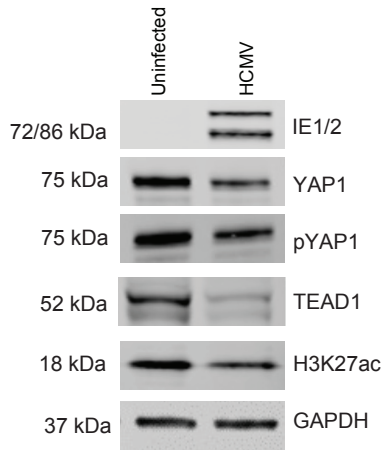
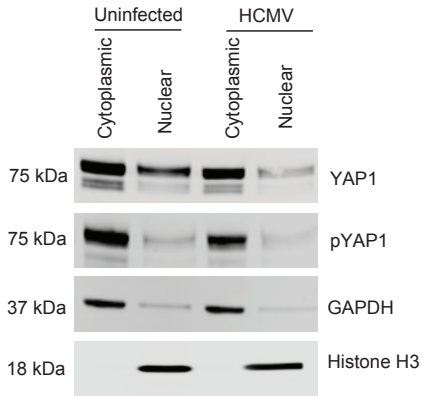


Figure 6

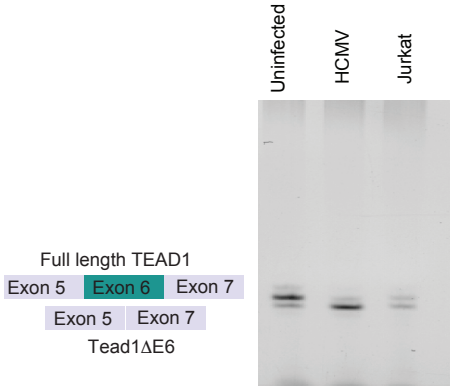
A. YAP1/TEAD1 protein expression levels



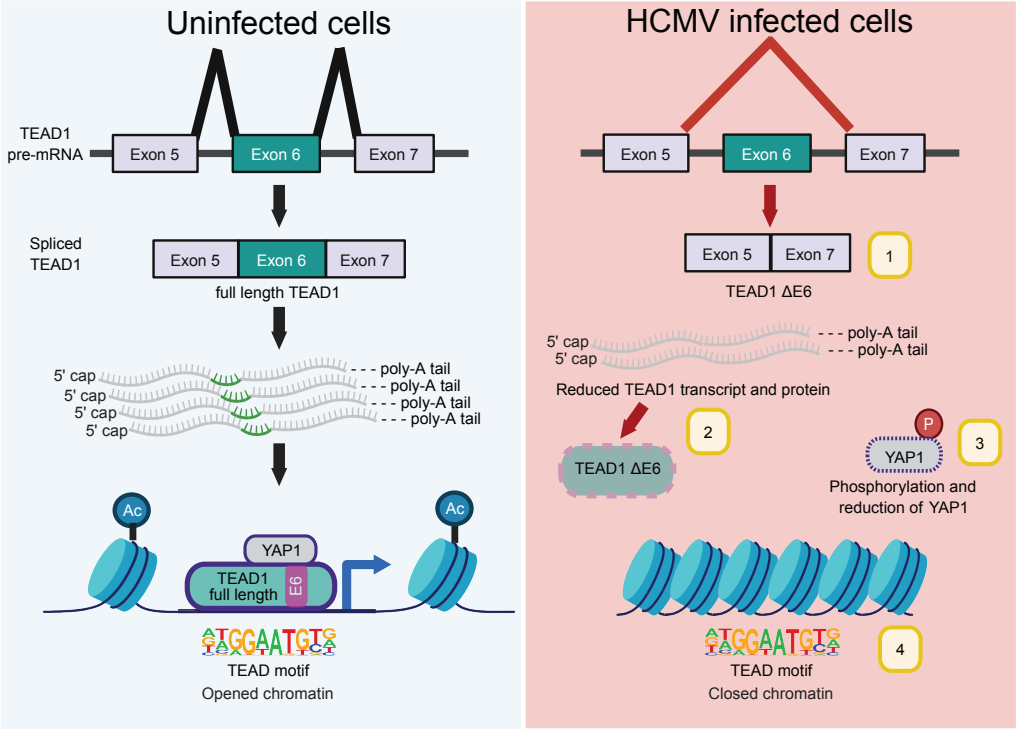
B. YAP1 protein localization



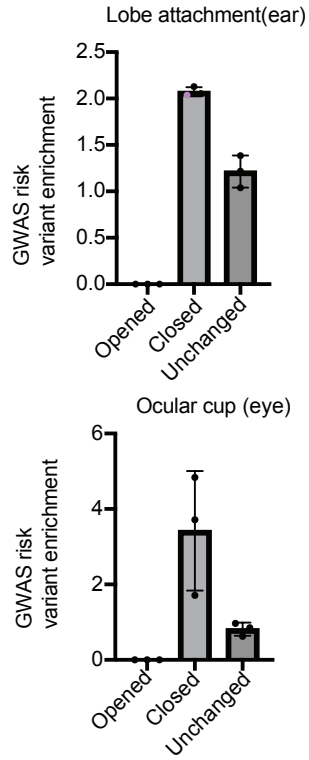
C. TEAD1 exon 6 alternative splicing validation



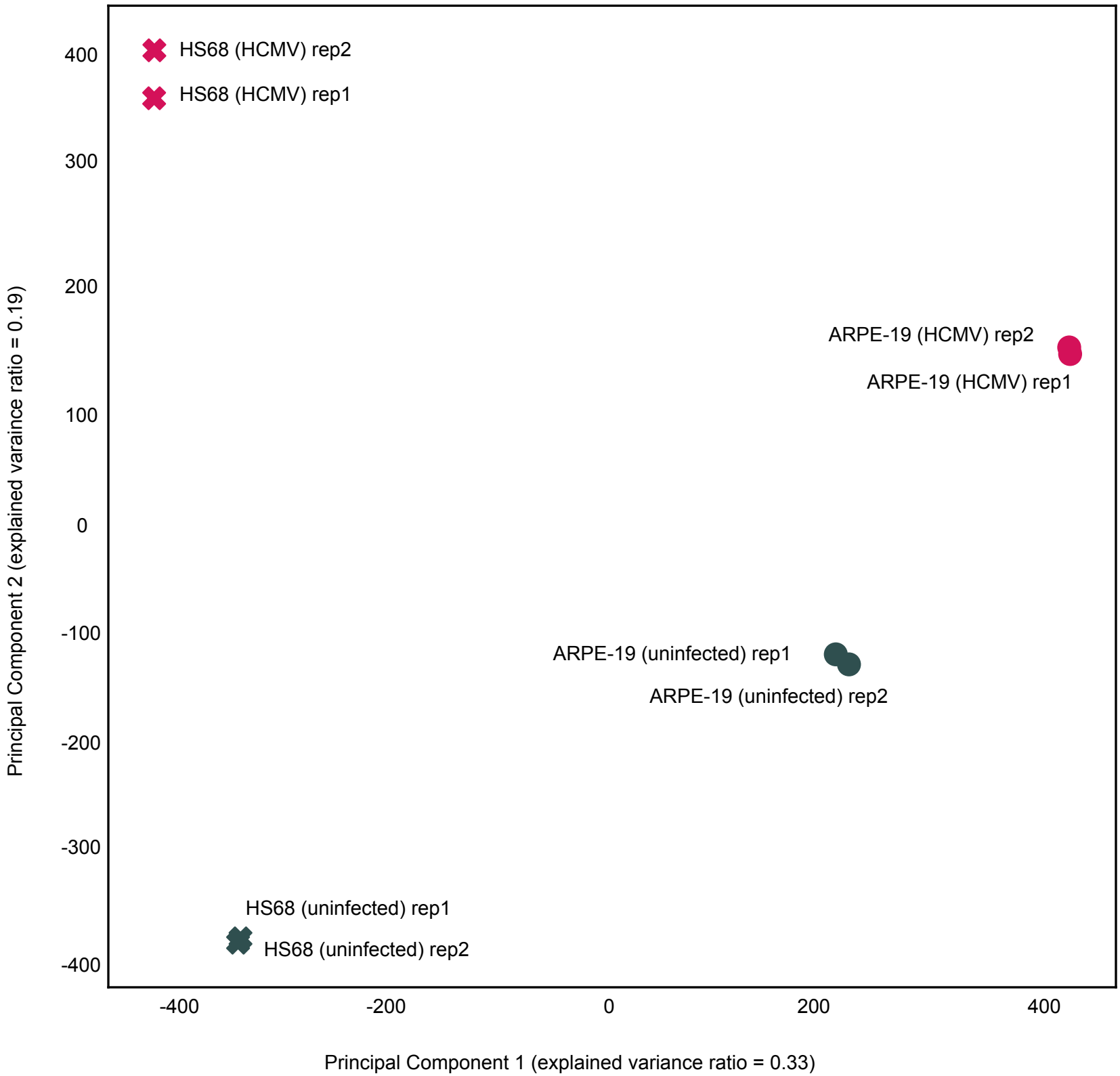
D. HCMV-induced mechanisms for downregulating TEAD1 activity



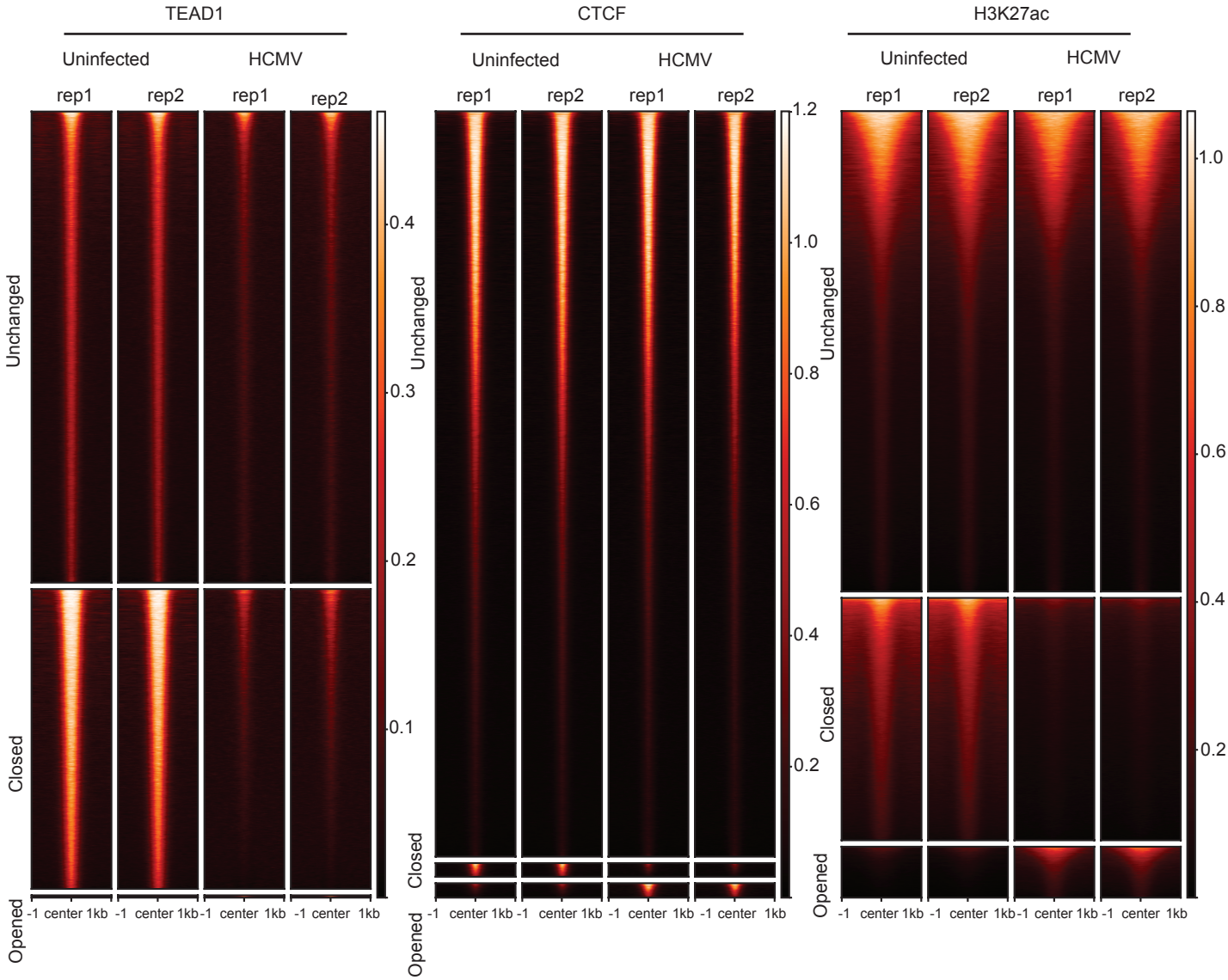
E. Enrichment of TEAD1 binding events at phenotype-associated genetic risk variants



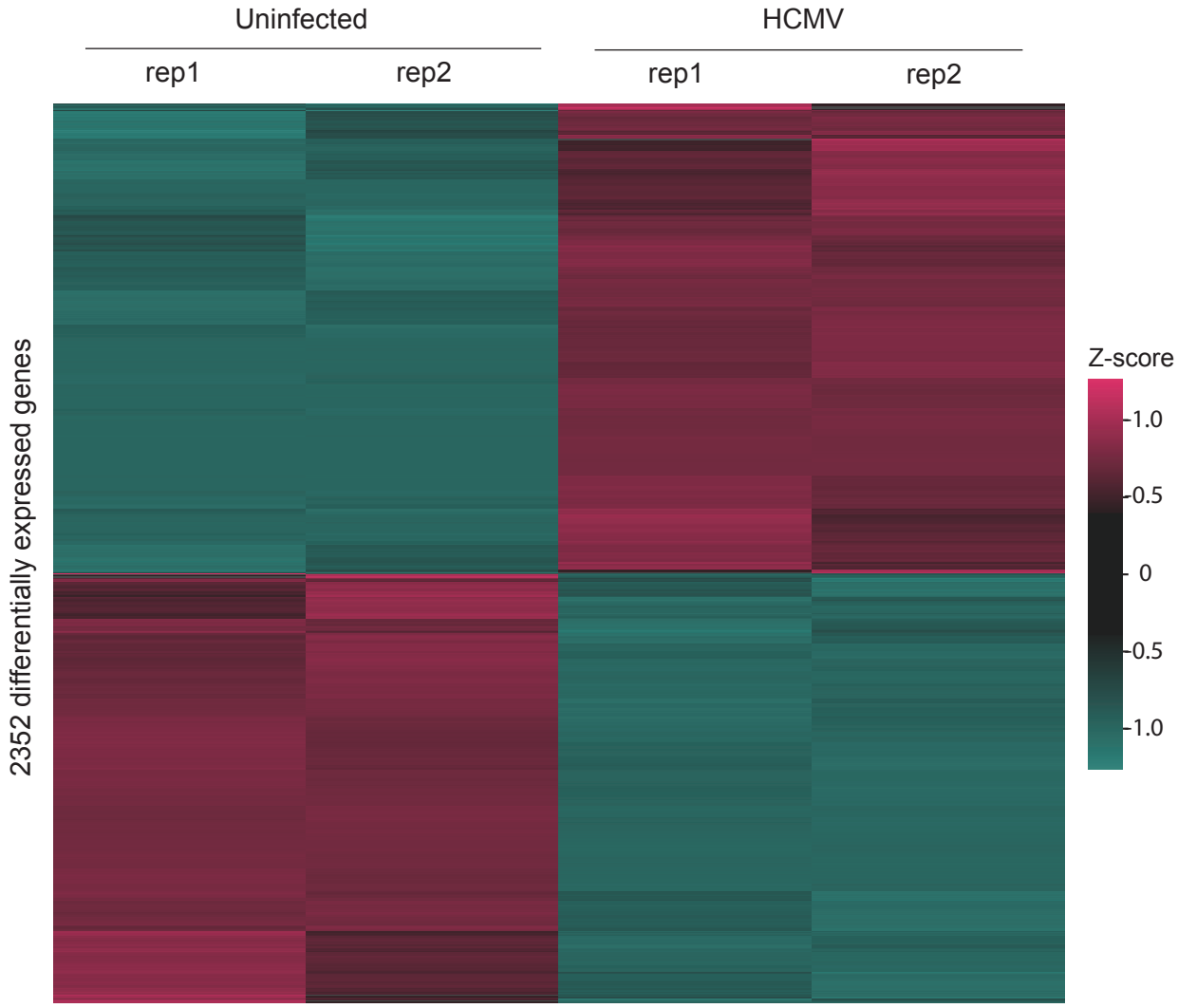
Supplemental Figure S1



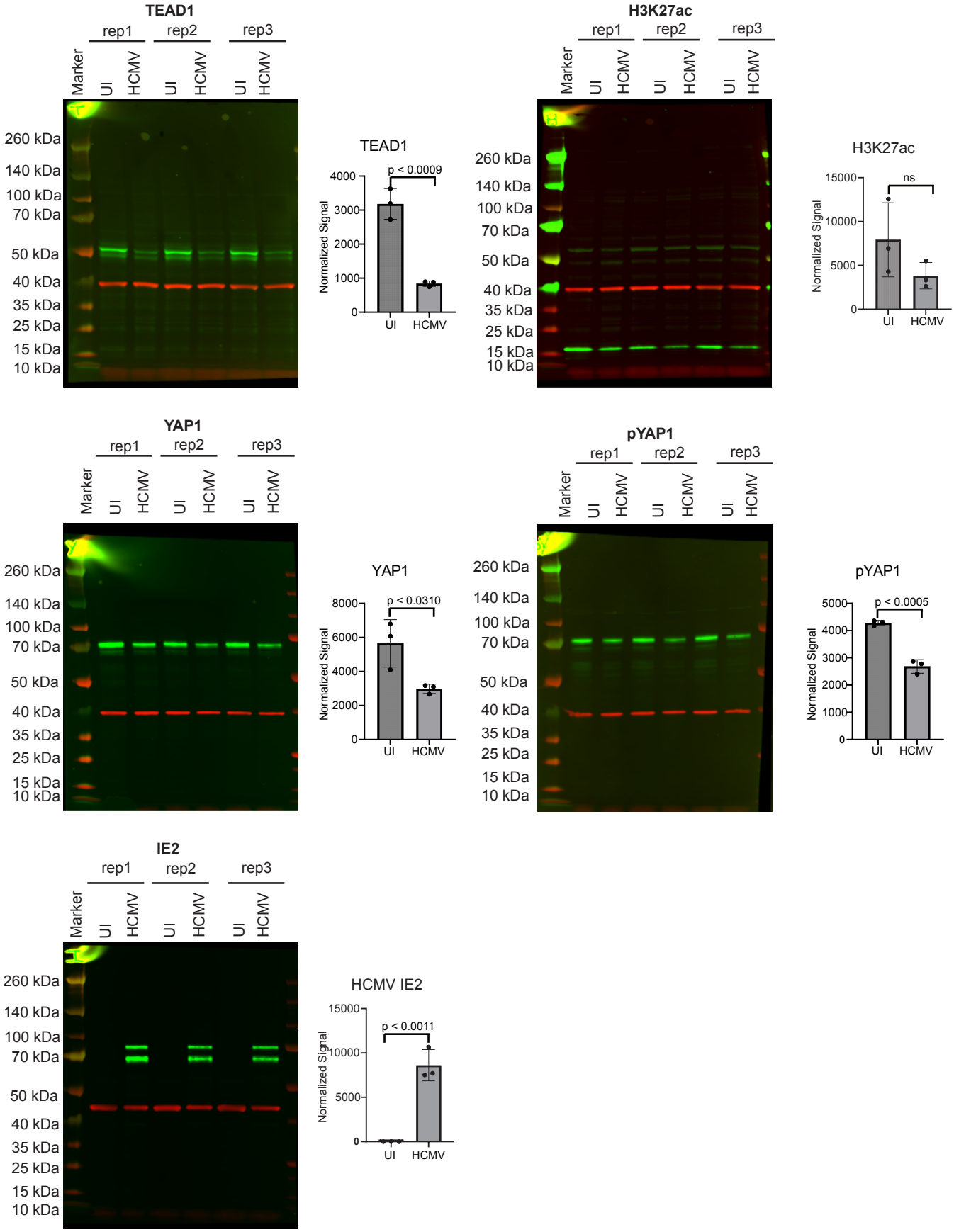
Supplemental Figure S2



Supplemental Figure S3



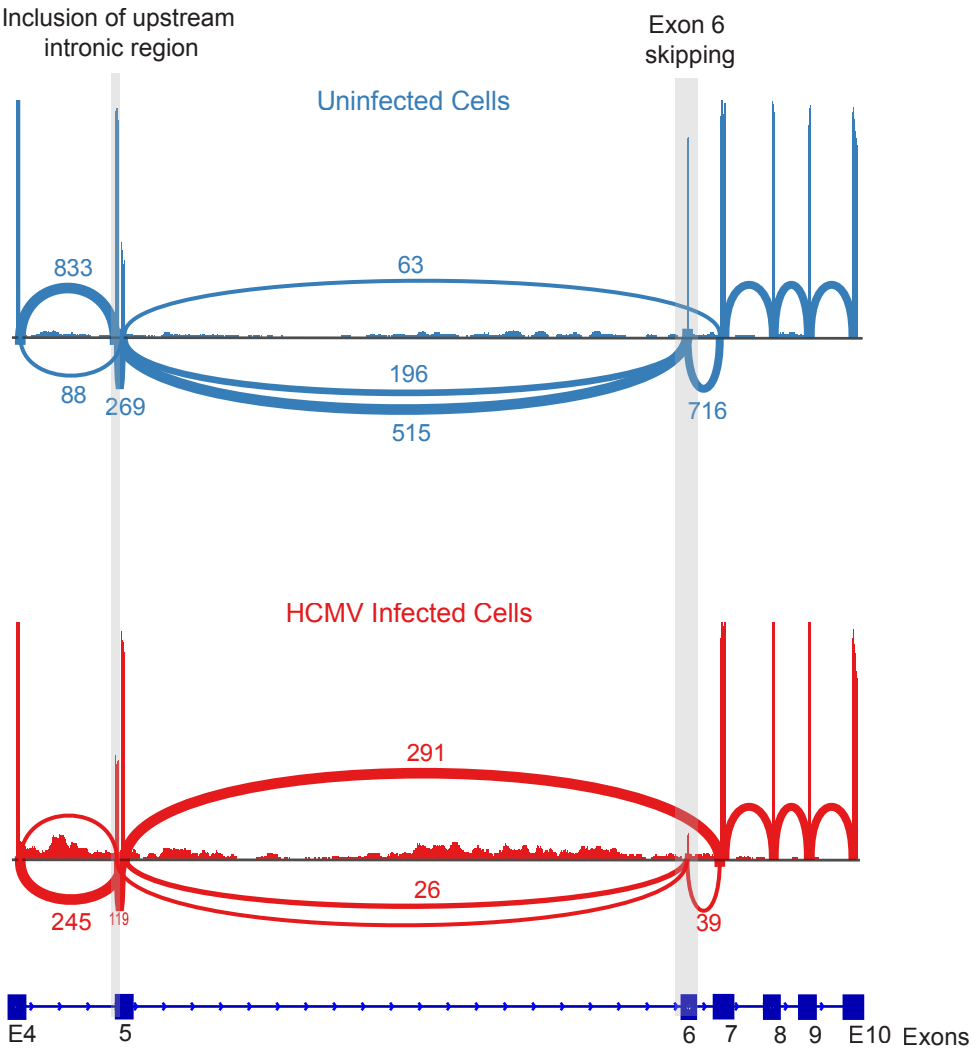
Supplemental Figure S4



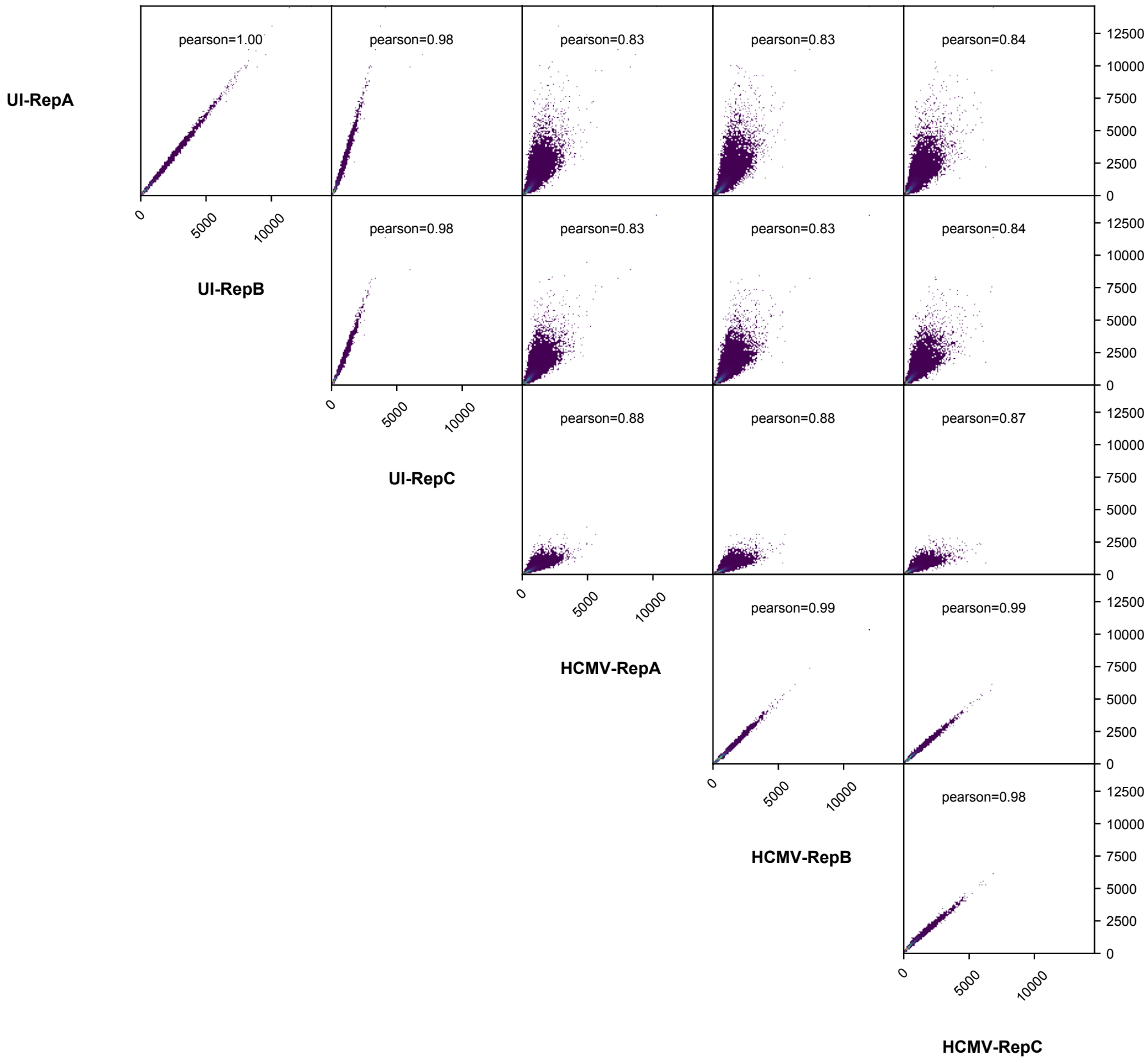
Supplemental Figure S5

A. Splicing counts

B. Sashimi plots



C. TEAD1 sequence showing PCR primers and exon 6 sequence

```
AAATCTCGTGATTTTCATTCCAAGCTAAAGGTAACAAGCATGGATCAGACTGCAAAGGATAAGGCCCTGCAGCACATGGCGGCCATGTCCTCAGCCCAGA
---------+---------+---------+---------+--------+---------+---------+---------+---------+---------+
TTTAGAGCACTAAAAGTAAGGTTCGATTTCCATTGTTCGTACCTAGTCTGACGTTTCCTATTCCGGGACGTCGTGTACCGCCGGTACAGGAGTCGGGTCT
K  S  R  D  F  H  S  K  L  K  V  T  S  M  D  Q  T  A  K  D  K  A  L  Q  H  M  A  A  M  S  S  A  Q  I
```

Supplemental Figure S6