HSV-1 infection induces a downstream-shift of the +1 nucleosome

Elena Weiß[1], Adam W. Whisnant[2,3], Thomas Hennig[2,3], Lara Djakovic[2], Lars Dölken[2,3,4#], Caroline C. Friedel[1#]

[1] Institute of Informatics, Ludwig-Maximilians-Universität München, Amalienstr. 17, 80333 Munich, Germany

[2] Institute for Virology and Immunobiology, Julius-Maximilians-University Würzburg, Versbacher Straße 7, 97078 Würzburg, Germany

[3] Institute for Virology, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, 30625 Hannover, Germany

[4] Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz-Center for Infection Research (HZI), 97080 Würzburg, Germany

[#] Corresponding authors

E-mail: lars.doelken@uni-wuerzburg.de; Caroline.Friedel@bio.ifi.lmu.de

Abstract

Herpes simplex virus 1 (HSV-1) infection induces a genome-wide loss of host transcriptional activity and widespread disruption of host transcription termination. The latter leads to read-through transcription for thousands of nucleotides beyond poly(A) sites and is associated with induction of open chromatin downstream of genes. In this study, we show that lytic HSV-1 infection alters chromatin accessibility around host promoters resulting in an extension of nucleosome-free regions at transcription start sites into downstream regions. This was most prominent for highly expressed genes and independent of the viral immediate-early proteins ICP0, ICP22, and ICP27 and the virion host shutoff protein *vhs*. ChIPmentation analysis of the noncanonical histone variant H2A.Z, which is strongly enriched at +1 and -1 nucleosomes, indicated that the downstream extension of accessible chromatin at promoters was linked to a downstream shift of +1 nucleosomes. In yeast, downstream shifts of +1 nucleosomes are induced by RNA Polymerase II (Pol II) degradation. Accordingly, treatment with phosphonoacetic acid (PAA), which inhibits viral DNA replication and alleviates the loss of Pol II from host genes, mostly abolished the downstream shift of +1 nucleosomes in HSV-1 infection. In contrast, irreversible inhibition of Pol II in human cells using α-amanitin similarly altered H2A.Z occupancy and +1 nucleosome positioning as HSV-1 infection. In summary, our study provides strong evidence that HSV-1-induced depletion of Pol II from the host genome leads to a widespread downstream shift of +1 nucleosomes at host gene promoters.

Author summary

Herpes simplex virus 1 (HSV-1) infection leads to a profound shutoff of host transcription during lytic infection. Loss of RNA Polymerase II (Pol II) in yeast has previously been shown to relax +1 nucleosome positioning to more thermodynamically favorable sites downstream of transcription start sites. In this study, we show that a similar phenomenon is likely also at play in lytic HSV-1 infection. Sequencing of accessible chromatin revealed a widening of nucleosome-free regions at host gene promoters into downstream regions. By mapping genome-wide positions of the noncanonical histone variant H2A.Z enriched at +1 and -1 nucleosomes, we demonstrate a downstream shift of +1 nucleosomes for the majority of cellular genes in lytic HSV-1 infection. We confirm that chemical inhibition of Pol II leads to a downstream shift of +1 nucleosomes in human cells. Thus, changes in chromatin architecture at promoters in HSV-1 infection are likely a consequence of HSV-1-induced depletion of Pol II activity from the host genome.

Introduction

Herpes simplex virus 1 (HSV-1) is one of nine human herpesviruses and more than half of the global population are latently infected with HSV-1 (1-4). HSV-1 does not only cause the common cold sores but is also responsible for life-threatening diseases particularly in the immunocompromised (5). The HSV-1 genome consists of 152 kilobases (kb) of double-stranded DNA and encodes for at least 121 large open reading frames (ORFs) and >100 small ORFs (6). Upon lytic infection, four of the five viral immediate early (IE) genes, i.e., ICP0, ICP4, ICP22 and ICP27, hijack the host gene expression machinery to ensure viral replication during early stages of infection (7-11). This results in recruitment of RNA polymerase II (Pol II) from the host chromatin to viral genomes and a general shutoff of host transcription (7, 9, 11-13). Pol II recruitment to viral genomes is predominantly facilitated by the viral IE protein ICP4 (13), which effectively sequesters Pol II into the viral replication compartments (14). However, ICP4 already depletes Pol II from cellular promoters as early as 2 h post infection (p.i.), i.e. before the onset of viral genome replication, and independent of viral genome copy number (15). The global host transcription shutoff is further exacerbated by ICP22-mediated inhibition of host transcription elongation (16) and the viral *vhs* (virion host shutoff) protein. *vhs* is a nuclease delivered by the tegument of the incoming viral particles that cleaves both viral and host mRNAs (17-19). We recently demonstrated that *vhs* continuously degrades about 30 % of host mRNAs per hour during the first 8 h of high multiplicity lytic infection of primary human fibroblasts (20).

HSV-1 further disturbs host transcription by inducing a widespread disruption of host transcription termination (DoTT) resulting in read-through transcription beyond polyadenylation (pA) sites of host genes for tens of thousands of nucleotides (nt) (21). Read-through transcription commonly extends into downstream genes, leading to a seeming

transcriptional induction of these genes. DoTT negatively affects host gene expression in two ways: First, as read-through transcripts are not exported (22), it effectively prevents translation of newly transcribed mRNAs and thus dampens the host transcriptional response to infection. Second, it redirects a large fraction of ongoing Pol II transcription from host gene bodies into downstream intergenic regions and, thus, further exacerbates the depletion of Pol II from host genes. We recently showed that ICP27 is sufficient for inducing DoTT (23). Interestingly, transcription downstream of genes (DoG) is also induced in cellular stress responses (22, 24-26). Accordingly, infection with an ICP27-null mutant still induces read-through transcription albeit at much lower levels, presumably due to a virus-induced cellular stress response (27). In contrast to stress-induced DoGs, HSV-1-induced DoTT is associated with a massive increase of chromatin accessibility downstream of genes (denoted as downstream open chromatin regions (dOCRs)) detectable by ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) (22). dOCR induction requires strong absolute levels of transcription downstream of genes, thus it is predominantly observed for highly transcribed genes with strong read-through. Treatment with phosphonoacetic acid (PAA) during HSV-1 infection, which inhibits viral DNA replication (28, 29), results in increased dOCR induction (30). As viral DNA redirects Pol II from the host chromatin, PAA treatment mitigates the loss of Pol II from the host genome and thus increases absolute levels of transcription on host genes and – importantly for dOCR induction – read-through regions. Although dOCRs can extend into downstream genes following read-through transcription, this occurs only in cases of very strong read-through transcription into these downstream genes. We recently showed that ICP22 is both required and sufficient for inducing dOCRs in presence of ICP27-induced read-through (30).

As we previously only investigated chromatin accessibility *downstream* of genes, we now further explore our previously published ATAC-seq data to investigate changes in chromatin

accessibility around host gene promoters. This revealed a broadening of open chromatin regions at the promoter into regions downstream of transcription start sites (TSS) for the majority of host genes. While it was consistently observed in absence of ICP0, ICP22, ICP27, and *vhs*, PAA treatment, which increases dOCR formation by alleviating the HSV-1-induced loss of transcription, substantially reduced the extension of open chromatin at the TSS. Weiner *et al.* previously showed that Pol II depletion in yeast induces downstream shifts of +1 nucleosomes, in particular for highly expressed genes (31). Transcribing Pol II disturbs histone-DNA interactions and at high transcription rates leads to eviction of nucleosomes from the chromatin (32). Weiner *et al.* proposed that Pol II depletion relaxes chromatin allowing nucleosomes to shift to more thermodynamically favorable sites. The noncanonical histone H2A variant H2A.Z is strongly enriched at gene promoters at +1 and -1 nucleosome positions (33, 34), likely as H2A.Z deposition increases nucleosome mobility and makes DNA more accessible to the transcriptional machinery (35). We thus performed ChIPmentation for H2A.Z during lytic HSV-1 infection to map +1 nucleosome positions. This showed that changes in chromatin accessibility in HSV-1 infection reflected a downstream shift of +1 nucleosomes. We observed similar changes in H2A.Z occupancy upon treatment with α-amanitin, a deadly toxin found in Amanita mushrooms, which inhibits transcription by preventing translocation of Pol II and triggering degradation of Rpb1, the largest subunit of Pol II (36). Our findings provide strong evidence that depletion of Pol II from host genes during HSV-1 infection leads to the downstream shift of +1 nucleosomes to more thermodynamically favorable sites.

Results

Widespread extension of chromatin-free regions downstream of transcription start sites

To investigate changes in chromatin accessibility within promoter regions during HSV-1 infection, we re-analyzed our recently published ATAC-seq data (30) for mock and wild-type (WT) HSV-1 infection of human fetal foreskin fibroblasts (HFF) as well as infection with null mutant viruses for the HSV-1 IE proteins ICP0, ICP22, and ICP27 as well as *vhs*. Samples for mock, WT, Δ*vhs* and ΔICP27 infection were collected at 8 h p.i., while samples for ΔICP0 and ΔICP22 infection were collected at 12 h p.i. as these null mutant infections progress more slowly. To assess chromatin accessibility changes in promoter regions, we identified genomic regions with differential ATAC-seq read density using our recently published RegCFinder method (37). RegCFinder searches for genomic subregions for which the read distribution differs between two conditions (e.g., mock and WT infection) and is targeted towards regions of interest by specifying genomic windows as input (**Sup. Fig 1a** in **S1 File**). For each input window, subregions are identified that show differences in the distribution of reads within this window. Statistical significance and log2 fold-changes for identified subregions are determined by DEXSeq (38). For our purposes, we defined the windows of interest as the ± 3kb promoter region around the TSS of 7,649 human genes. These TSSs were previously defined in our analysis of promoter-proximal Pol II pausing in HSV-1 WT infection (39) based on a re-analysis of published precision nuclear run-on sequencing (PRO-seq) and PROcap-seq (a variation of PRO-seq) data of flavopiridol-treated uninfected HFF. CDK9 inhibition by flavopiridol arrests Pol II in a paused state at the TSS, allowing the mapping of Pol II initiation sites with both PRO-seq and PROcap-seq. Filtering for high-confidence TSSs that were confirmed by both data sets and within a maximum distance of 500 bp to the nearest annotated gene on the same strand identified the major TSS for 7,649 genes. Using these 6 kb promoter windows as input, we applied RegCFinder to our ATAC-seq data for WT and all

null-mutant infections in comparison to mock to identify promoter subregions that show differences in the distribution of open chromatin between infection and mock. We did not filter these promoter windows beforehand to exclude genes for which dOCRs extended from an upstream gene into their promoter window for two reasons: (i) Since dOCRs represent changes in chromatin accessibility, RegCFinder should by design detect them; (ii) The further the distance from the upstream gene, the less pronounced are changes in chromatin accessibility in dOCR regions, making it difficult to distinguish them from background. Any approach we tested to filter promoter windows overlapping with dOCRs from upstream genes before the RegCFinder analysis either excluded too many or too few windows. Thus, we decided to first run RegCFinder and then investigate which of the detected changes were due to induction of dOCRs for upstream genes.

RegCFinder identified 23,000-24,000 differential subregions (in the following also denoted as RegC, short for regions of change) for each virus infection compared to mock (WT: 23,357 RegC, ΔICP0: 23,532, ΔICP22: 24,640, ΔICP27: 23,523, Δ*vhs*: 23,716, **Sup. Fig 1b** in **S1 File**). Between 29 and 37 % of these showed a statistically significant difference in read density within the corresponding 6 kb promoter windows upon infection (multiple testing adjusted p-value (adj. p.) ≤ 0.01, WT: 8,552 RegC, ΔICP0: 6,740, ΔICP22: 5,710, ΔICP27: 7,188, Δ*vhs*: 6,860). These represented between 3,129 (ΔICP22 infection) and 4,391 (WT) genes (**Sup. Fig 1c** in **S1 File**). The lower fraction of statistically significant differential regions in ΔICP22 infection is likely due to the relatively low number of ATAC-seq reads mapping to the host genome compared to the other infections (~3.8-fold fewer, see Supplementary Data 1 in (30)). **Fig 1a** visualizes the positions of identified RegC for the 4,981 promoter windows containing at least one statistically significant (adj. p. ≤ 0.01) differential region for at least one virus infection compared to mock and **Sup. Fig 1d** in **S1**

File illustrates log2 fold-changes for statistically significant regions. Here, blue indicates that relative read density in that subregion is increased during virus infection compared to mock (infection>mock, denoted as i-RegC in the following) and red that relative read density in that subregion is higher in mock compared to infection (mock>infection, denoted as m-RegC).

Strikingly, our results showed a highly homogeneous picture for WT and null-mutant infections compared to mock, with changes in the same direction observed at approximately the same locations. To identify distinct patterns of changes in chromatin accessibility, we performed hierarchical clustering on the RegC location heatmap in **Fig 1a** and selected a cutoff on the dendrogram to obtain 14 clusters identified by visual inspection. To visualize average read density for each cluster, we performed metagene analyses of promoter windows combined with statistics on RegC locations for each cluster (**Fig 1b,d,f,g, Sup. Fig 2** in **S1 File**). For metagene analyses, each 6 kb input window was divided into 101 bins of ~59 bp. For each bin, ATAC-seq read counts were determined, normalized to sequencing depth, normalized to sum up to 1 for each input window, averaged across all windows in each cluster and then averaged across replicates. In addition, we show which fraction of windows have an m- (red) or i-RegC (blue) for each bin. Based on the location heatmap and the metagene analyses, we identified three major patterns of changes in chromatin accessibility around promoters during WT infection: (i) ATAC-seq peaks at the TSS that shift and/or broaden into regions downstream of the TSS during HSV-1 infection (clusters 1, 2, 4, 5, 9, 10, 12, 14, a total of 3,472 genes, **Fig 1b, Sup. Fig 2a-g** in **S1 File,** examples in **Fig 1c, Sup. Fig 3a-h** in **S1 File**), (ii) ATAC-seq peaks at the TSS that shift and/or broaden into regions upstream of the TSS (clusters 6, 7, 11, a total of 919 genes, **Fig 1d, Sup. Fig 2h,i** in **S1 File,** examples in **Fig 1e, Sup. Fig 3i-k** in **S1 File**), and (iii) an increase in chromatin-accessibility both up- and downstream of the TSS peak in infection compared to mock (cluster 13, 126 genes, **Fig 1f,** example in **Sup. Fig 3l** in **S1 File**). Two clusters (3 and 8, a total of 464 genes) exhibited a

combination of patterns (i) and (ii) with an extension of the TSS peak in both up- and downstream direction (**Fig 1g, Sup. Fig 2j, Sup. Fig 3m,n** in **S1 File**).

Most chromatin accessibility changes at promoters are independent of dOCR induction

First, we evaluated which of the observed changes were due to dOCRs extending into promoter regions. For this purpose, we determined the fraction of promoter windows in each cluster that overlapped with dOCRs of the 1,296 genes for which we previously showed consistent dOCR induction across different HSV-1 strains (30) (**Fig 2a**). dOCR regions in mock, WT and all null mutant infections (average across two replicates) were identified as previously described (22, 30) (see also Materials and Methods). It is important to note here that some relatively short dOCRs are aready observed in mock infection but these substantially extend during HSV-1 infection for read-through regions. In this analysis, we also included ATAC-seq data obtained in the same experiment for 8 h p.i. WT infection combined with PAA treatment (WT+PAA) to identify dOCR regions that are not as clearly detectable without PAA treatment but may still bias promoter analyses. WT (± PAA) and null mutant infections are ordered according to the overall extent of dOCR induction in **Fig 2a**. While no differences to mock infection were observed for ΔICP22 infection as expected, some clusters showed an enrichment for dOCR overlaps in other virus infections. This increased with the overall extent of dOCR induction in these viruses. However, no cluster had more than 20% of promoter windows overlapping with dOCRs even with PAA treatment and most had <5% overlap, indicating that dOCRs represent only a very small fraction of observed changes in chromatin accessibility around promoters. Most importantly, the extension of accessible chromatin regions at promoters was also observed upon infection with an ICP22-null mutant indicating that the extension is not linked to dOCR formation (**Sup. Fig 4-17** in **S1 File**). As exemplified for cluster 5 (pattern (i)) and cluster 7 (pattern (ii)) in **Fig 2b** and **c**, respectively, the extension was still detected in absence of ICP22 but strongly reduced by PAA treatment.

Moreover, even for clusters 12-14, which exhibited some enrichment for dOCR overlaps, most differential regions were also detected in ΔICP22 infection (**Sup. Fig 15-17** in **S1 File**) and thus not linked to dOCR induction. As absence of ICP0, ICP27 or *vhs* also had little impact on changes in chromatin accessibility around promoters, we conclude that neither dOCR induction nor activity of ICP0, ICP22, ICP27 or *vhs* alone explain the majority of observed changes in chromatin accessibility at promoters in HSV-1 infection.

Extension of accessible chromatin regions around promoters is linked to transcription

To correlate observed changes to potential transcription factor binding, we next performed motif discovery for novel and known motifs in m- and i-RegC grouped either by cluster or pattern compared to the background of all promoter windows using HOMER (40), but found no enriched motifs. Functional enrichment analysis for Gene Ontology (GO) terms for each cluster also yielded no significant results, with the exception of cluster 13 (pattern (iii)). Cluster 13 was enriched for "mRNA splicing via the spliceosome" and related terms, however, this was mainly due to several snRNA genes included in this cluster. Accordingly, cluster 13 was significantly enriched for snRNA genes (**Fig 2d**, adj. p. = 3×10^{-16}, see Materials and Methods). In contrast, pattern (i) clusters tended to have relatively high fractions of protein-coding genes. Strikingly, cluster 7 (= the most pronounced case of pattern (ii)) was significantly enriched for antisense transcripts (adj. p. = 5×10^{-15}). The other pattern (ii) clusters 6 and 11 as well as cluster 8 (combined pattern (i) + (ii)) also contained a relatively high fraction of antisense transcripts, although this was not statistically significant. Consistently, we found that cluster 7 (adj. p. = 3.8×10^{-14}) and to a lesser degree clusters 6 and 11 (adj. p. = 0.006 and 0.041, respectively) were enriched for bidirectional promoters, i.e. promoters containing an annotated gene start within 1 kb upstream of the TSS of the target gene (= the gene around whose TSS the promoter window was originally defined, **Fig 2e,** see **Fig 1e** for an example).

We thus hypothesized that pattern (ii) essentially represented the mirror image of pattern (i) (mirrored at a vertical axis through the TSS) with the “downstream” broadening/shift of chromatin accessibility occurring in antisense direction for bidirectional promoters. To confirm this, we calculated the ratio of sense to antisense transcriptional activity in promoter windows for all genes. For this purpose, we analyzed RNA-seq data of chromatin-associated RNA, which depicts nascent transcription, for mock and 8 h p.i. WT HSV-1 infection from our recent study (20). We also included in this analysis the 2,668 genes (denoted as NA group) without significant chromatin accessibility changes that were excluded from **Fig 1a**. Interestingly, metagene analyses for these genes also revealed a slight broadening of the TSS peak into downstream regions, although this was much less pronounced than for pattern (i) clusters and therefore not detected by RegCFinder (**Sup. Fig 18a** in **S1 File**). To calculate the ratio of sense to antisense transcription, promoter windows were divided into the regions down- and upstream of the TSS and expression in chromatin-associated RNA was determined in sense direction for the downstream region (=DSR) and in antisense direction for the upstream region (=UAR). We then compared log2 ratios of DSR to UAR expression between clusters (**Fig 3a)**. Here, positive log2(DSR:UAR) ratios indicate that sense transcription downstream of the TSS is stronger than antisense transcription upstream of the TSS and negative values indicate the opposite. Strikingly, all pattern (ii) clusters had significantly lower DSR to UAR ratios than the remaining genes, while pattern (i) tended to have significantly higher ratios. While for clusters 6 and 11 with pattern (ii) median values were still positive, values for cluster 7 (most pronounced pattern (ii)) were commonly negative. Thus, transcription upstream of the TSS in antisense direction was the dominant mode of transcription for cluster 7 promoters and pattern (ii) essentially just represents pattern (i) for the genes on the antisense strand.

While we observed no consistent differences in gene expression changes in WT compared to mock infection between patterns (**Sup. Fig 18b** in **S1 File**), gene expression (quantified as FPKM = fragments per kilobase million mapped reads, **Fig 3b**) differed considerably between patterns. Most pattern (i) clusters as well as cluster 3 (combined pattern (i)+(ii)) exhibited median expression levels above average in mock (**Fig 3b**), which was statistically significant for 5 clusters. In contrast, clusters 6, 7 (pattern (ii)) and 8 (combined pattern (i)+(ii)) exhibited relatively low expression values, with cluster 7 genes having significantly lower expression. The latter is consistent with these genes being less expressed than their antisense counterpart in these bidirectional promoters. Genes without significant chromatin accessibility changes (NA group) also showed significantly lower expression, albeit not as low as cluster 7. These differences between clusters/patterns were generally maintained in HSV-1 infection at lower overall expression levels, consistent with the absence of differences in gene expression fold-changes between most clusters. In summary, these results indicate that HSV-1 infection leads to an extension of the accessible chromatin around promoters in the dominant direction of transcription for the majority of host genes, with highly expressed genes being most strongly affected.

Consistent with transcription being linked to changes in chromatin accessibility, PAA treatment, which alleviates the depletion of Pol II from host genomes, substantially reduced the broadening and/or downstream shift of accessible chromatin regions at promoters (**Fig 2b,c, Sup. Fig 4-17** in **S1 File**) as well as the fold-changes compared to mock for differential regions (**Sup. Fig 18c** in **S1 File**). Exceptions were cluster 13, and parts of cluster 12 and 14, for which chromatin accessibility increased up- (for clusters 12-14) and downstream (for cluster 13) of the TSS peak (**Sup. Fig 15-17** in **S1 File**). Since these clusters were enriched for overlaps to dOCRs from an upstream gene, these changes reflect increased dOCR induction

upon PAA treatment. Notably, however, the broadening of the TSS peak into downstream regions in clusters 12 and 14 was greatly reduced by PAA treatment.

To investigate how early in infection changes in chromatin accessibility around promoters can be detected, we ran RegCFinder on an ATAC-seq time-course for 1, 2, 4, 6, and 8 h p.i. WT infection from our previous publication (22) (**Fig 3c,d**). While barely any significant differential regions were identified at 1 and 2 h p.i., 1,963 significant regions were identified in 1,249 promoter windows by 4 h p.i. and 7,558 significant regions in 3,989 promoter windows by 8 h p.i. (**Fig 3d**). Log2 fold-changes for significant differential regions generally reflected the patterns observed in the analysis for WT and null mutant infections (**Fig 3c**). The same applied when we separately determined fold-changes and significance on the time-course ATAC-seq data for the differential regions determined for WT vs. mock from our initial analysis (**Sup. Fig 19** in **S1 File**) and when we performed metagene analyses for the clusters from **Fig 1a (Fig 3e**, **Sup. Fig 20** in **S1 File**, example genes in **Sup. Fig 21** in **S1 File)**. Although the changes in chromatin accessibility at 8 h p.i. of the time-course were less pronounced than for the original WT vs. mock comparison at 8 h p.i., this nevertheless confirms the changes in chromatin accessibility around promoters in HSV-1 infection in an independent experiment. Furthermore, it reveals that these changes begin to manifest between 2 and 4 h p.i. We previously showed that transcriptional activity on host genes drops substantially in the first 4 h of infection to only 40% of transcription in uninfected cells, which was further halved until 8 h p.i. (20). Thus, the onset of the extension of accessible chromatin regions at promoters into downstream (for pattern (i)) and upstream (for pattern (ii)) regions coincides with the onset of the drop in host transcription during infection.

HSV-1 infection induces a downstream shift of +1 nucleosomes

The HSV-1-induced changes in chromatin accessibility at promoters suggest a broadening of the nucleosome-free region around promoters in either sense (pattern (i)) or antisense (pattern (ii)) direction. Consistent with this, we observed downstream and upstream shifts for pattern (i) and (ii) clusters, respectively, in the nucleosome-depleted regions at promoters in H3K36me3 ChIPmentation data for mock and 8 h WT HSV-1 infection from our previous study (30) (**Sup. Fig 22** in **S1 File**). H3K36me3 marks H3 histones on actively transcribed genes (41). These changes in nucleosome-free regions require a shift in nucleosome positioning, in particular for +1 or -1 nucleosomes. Several studies performed in yeast previously investigated the role of transcriptional activity on nucleosome positioning (42-44). In particular, Weiner *et al.* reported a downstream shift of +1 nucleosomes upon Pol II depletion in yeast (31). Here, the strongest downstream shifts were observed for very highly expressed genes, consistent with our observation that pattern (i) was associated with higher gene expression (**Fig 3b**). To investigate whether changes in chromatin accessibility at promoters during HSV-1 infection were linked to shifts in +1 and -1 nucleosome positions, we performed ChIPmentation in HFF for mock and WT infection at 8 h p.i. with an antibody recognizing the C-terminal part of the non-canonical histone H2A.Z (n=3 replicates). H2A.Z is highly enriched at gene promoters at -1 and +1 nucleosome positions (33, 34) and is encoded by two genes, whose protein products H2A.Z.1 and H2A.Z.2 differ by only three amino acids. While their genomic occupancy patterns are similar, there are quantitative differences with H2A.Z.1 being more abundant at active promoters than H2A.Z.2 (45). Since the C-terminal regions of H2A.Z.1 and H2A.Z.2 differ only by the last amino acid, the antibody used for ChIPmentation recognizes both isoforms. A metagene analysis of all analyzed promoter windows showed the expected distribution of H2A.Z occupancy in mock infection with two peaks on both sides of the TSS, corresponding to -1 and +1 nucleosome positions (**Sup. Fig 23a** in **S1 File**).

Application of RegCFinder to H2A.Z ChIPmentation data generally identified a relative increase in H2A.Z downstream of the TSS during infection and a relative decrease upstream of the TSS for pattern (i) and combined pattern (i)+(ii) clusters (**Fig 4a**, log2 fold-changes for differential regions shown in **Sup. Fig 23b** in **S1 File**, examples in **Fig 4b**, **Sup. Fig 24a-g,k,l** in **S1 File**). Metagene analyses showed that this reflected a downstream shift and broadening of +1 nucleosome peaks as well as a relative increase of +1 nucleosome peaks compared to -1 nucleosome peaks (**Fig 4d**, **Sup. Fig 23c-i,m,n** in **S1 File**). In contrast, cluster 7 (most pronounced pattern (ii) cluster) showed the opposite trend with relative increases in H2A.Z upstream of the TSS and decreases downstream of the TSS (example in **Fig 4c**). Consistent with pattern (ii) representing pattern (i) in antisense direction, the metagene analysis showed an upstream shift and broadening of the -1 nucleosome peak and a relative increase of the -1 nucleosome peak compared to the +1 nucleosome peak (**Fig 4e**). However, this was only observed for few genes in cluster 6 and cluster 11 (**Fig 4a**, **Sup. Fig 23j,k** in **S1 File**, examples in **Sup. Fig 24h,i** in **S1 File**), which fits with the observation that pattern (ii) in the ATAC-seq data was also much less pronounced for these clusters. In contrast to pattern (i) and (ii), pattern (iii) in the ATAC-seq data (cluster 13) was not associated with distribution changes in H2A.Z (**Sup. Fig 23l**, example in **Sup. Fig 24j** in **S1 File**), indicating that it is not shaped by shifts in +1 and/or -1 nucleosome positioning. This is consistent with pattern (iii) being at least partly associated with dOCR induction for upstream genes, which we previously linked to impaired nucleosome repositioning following Pol II transcription downstream of genes (30).

In summary, our data suggest that the alterations in chromatin accessibility around promoters result from an extension of the nucleosome-free region around promoters due to downstream and upstream shifts of +1 or -1 nucleosomes, respectively. Here, the direction of shift depends on whether sense or antisense transcription represents the dominant direction of transcription

in this promoter. Notably, a recent study in yeast proposed that -1 nucleosomes should actually be considered +1 nucleosomes for antisense transcription (46). We hypothesized that – similar to what was previously observed for yeast – the loss of Pol II from host promoters during HSV-1 infection induced the shift of +1 and -1 nucleosome positions. To confirm that inhibition of transcription alone can lead to similar shifts of +1 and -1 nucleosomes in human cells, we re-analyzed published H2A.Z ChIP-seq data of HCT116 cells with and without α-amanitin treatment from the recent study by Lashgari *et al.* (47). They showed that α-amanitin treatment leads to reduced Pol II levels at gene TSSs and increased incorporation of H2A.Z at TSSs of transcribed genes. However, they only analyzed changes in absolute H2A.Z levels at promoters and did not focus on changes in +1 and -1 nucleosome positioning. Metagene and RegCFinder analyses of the H2A.Z ChIP-seq data indeed showed a similar trend upon α-amanitin treatment as during HSV-1 infection (**Fig 4f,g, Sup. Fig 25,** examples are shown in **Sup. Fig 26** in **S1 File**). Pattern (i) and pattern (i)+(ii) clusters showed a massive broadening of the +1 nucleosome peak into downstream regions, even more pronounced than what is observed upon HSV-1 infection. In addition, they also showed a less pronounced broadening of -1 nucleosome peaks. Similarly, pattern (ii) clusters showed a strong broadening of both -1 and +1 nucleosome peaks into upstream regions. In case of cluster 7 (strongest pattern (ii)), the broadening was much more pronounced for the -1 nucleosome (i.e. the +1 nucleosome in antisense direction) than for the +1 nucleosome (i.e. the -1 nucleosome in antisense direction). This analysis confirms that depletion of Pol II from promoters leads to a downstream shift of +1 nucleosomes in human cells. Together with the mitigating effect of PAA treatment on corresponding chromatin accessibility changes in HSV-1 infection and the timing of onset, our results thus provide strong evidence that the shift of +1 and -1 nucleosomes in the direction of transcription in HSV-1 infection is a consequence of Poll II depletion from host chromatin.

Discussion

Previously, we showed that HSV-1 infection disrupts chromatin architecture downstream of genes with strong read-through transcription. In this study, we reveal that chromatin architecture is also substantially altered at host gene promoters during HSV-1 infection. Here, 57% of genes showed a statistically significant change in the distribution of chromatin accessibility at promoters in WT HSV-1 infection and metagene analyses indicated similar but less pronounced changes even for genes without statistically significant changes. In essence, we identified three types of changes: Most genes showed a shift and/or broadening of accessible chromatin into regions downstream of the TSS (pattern (i)), while ~900 genes showed the opposite trend with a shift and/or broadening of accessible chromatin into regions upstream of the TSS (pattern (ii)). A further ~460 genes exhibited a combination of pattern (i) and (ii), with the broadening of accessible chromatin in downstream direction being more pronounced than in upstream direction. Only a small set of genes (126 genes) showed relative increases in chromatin accessibility up- and downstream of the TSS, which could partly be attributed to overlaps with dOCRs from upstream genes and partly to an enrichment for (short) non-coding RNAs for which promoter windows were larger than the actual gene length. We thus did not further investigate this pattern.

As patterns (i) and (ii) were still observed in absence of ICP22, which is necessary for dOCR induction, these are not artefacts of dOCRs extending into downstream genes. On the contrary, PAA treatment, which increases dOCR induction, substantially reduced – though not completely abolished – pattern (i) and (ii). Furthermore, knockdown of neither ICP0, ICP22, ICP27 nor *vhs* substantially affected pattern (i) and (ii). This led us to hypothesize that the loss of Pol II from host genes during HSV-1 infection, which is alleviated by PAA, causes

the widespread extension of open chromatin at host gene promoters. Analysis of transcriptional activity using RNA-seq of chromatin-associated RNA revealed that pattern (i) and (ii) are indeed linked to transcription, with pattern (i) associated with more highly expressed genes and pattern (ii) associated with bidirectional promoters with strong antisense transcription on the opposite strand. In case of cluster 7, for which pattern (ii) was most pronounced, antisense transcription was indeed much higher than sense transcription. This was not due to the widespread induction of antisense transcription in HSV-1 infection, which we previously reported on (48), but these genes already exhibited strong transcription in antisense direction in uninfected cells originating from bidirectional promoters. Consistently, cluster 7 was enriched for genes annotated as antisense and more than a third of these promoters contained an annotated gene starting on the opposite strand within 1 kb upstream of the TSS of the target gene. Thus, pattern (ii) is essentially just pattern (i) for the genes on the opposite strand in bidirectional promoters.

Since the broadening of chromatin accessibility in promoter regions suggested an extension of nucleosome-free regions at promoters, we mapped +1 and -1 nucleosome positions by ChIPmentation of the H2A.Z histone variant enriched at these nucleosomes. Analysis of H2A.Z occupancy indeed showed a downstream shift of +1 nucleosomes for pattern (i) genes in HSV-1 infection and an upstream shift of -1 nucleosomes for genes with most pronounced pattern (ii) (cluster 7). This further confirmed that pattern (ii) just represents pattern (i) for genes on the antisense strand. It is also consistent with recent results from yeast by Bagchi *et al.* that -1 H2A.Z-containing nucleosomes should actually be considered as +1 nucleosomes for antisense transcription (46). Downstream shifts of +1 nucleosomes have previously been reported upon Pol II degradation in yeast by Weiner *et al.* (31), in particular for highly expressed genes. Re-analysis of published H2A.Z ChIP-seq data with and without α-amanitin-induced depletion of Pol II from gene promoters confirmed that loss of Pol II alone is

sufficient to induce downstream shifts of +1 nucleosomes in human cells. Moreover, for genes with dominant antisense transcription (cluster 7), it resulted in upstream shifts of -1 nucleosomes. Here, the broadening of H2A.Z peaks in both HSV-1 infection and upon α-amanitin treatment, respectively, indicate a less precise positioning of corresponding nucleosomes. This was more pronounced upon α-amanitin treatment and also observed for -1 nucleosomes, suggesting that loss of Pol II upon α-amanitin treatment is more pronounced than in HSV-1 infection.

The +1 nucleosome has been shown to play a role in promoter-proximal Pol II pausing, which occurs between the promoter and the +1 nucleosome (49). It is thus tempting to speculate that the downstream shifts in +1 nucleosomes we observed in this study may provide an explanation for the downstream shift in Pol II pausing in HSV-1 infection we recently reported (39). Notably, there are a few parallels between the downstream shifts of +1 nucleosomes and changes in Pol II pausing during HSV-1 infection. Both appear to be established relatively early in infection, with the extension of the nucleosome-free region detectable at 4 h p.i. but not at 2 h p.i. in our ATAC-seq data and the downstream shift of Pol II pausing clearly detectable in PRO-seq data at 3 h p.i. but not yet at 1.5 h p.i. (49). Furthermore, both +1 nucleosomes and Pol II pausing sites, which are well-positioned in uninfected cells, appear much less well-positioned in HSV-1 infection. In this way, both phenomena would largely be a bystander effect of the loss of host transcriptional activity and not necessarily play a functional role for HSV-1 infection. Alternatively, the shifts in +1 nucleosomes might also be linked to the mobilization of H1, H2 (including H2A.Z) and H4 histones during HSV-1 infection (50-52), which has been proposed to serve as a source of histones for viral chromatin assembly. In any case, our study highlights that HSV-1 infection impacts chromatin architecture at promoters independently of the widespread changes downstream of genes mediated by ICP22.

Materials and Methods

Previously published sequencing data analyzed in this study

ATAC-seq for mock and WT HSV-1 infection (strain 17, 8 h p.i.), WT infection with 8 h PAA treatment, infection with ICP0-null mutant (ΔICP0, strain 17, (53), 12 h p.i.), ICP22-null mutant (ΔICP22, R325, strain F, (54), 12 h p.i.), ICP27-null mutant (ΔICP27, strain KOS, (55), 8 h p.i.) and *vhs*-null mutant (Δvhs, strain 17, (56)) in of HFF were taken from our recent publication (30) (n=2 apart from ΔICP22 infection with n=4, GEO accession: GSE185234). ATAC-seq data for mock and WT HSV-1 infection of HFF at 1, 2, 4, 6 and 8 h p.i. WT infection (n=2 replicates, GEO accession: GSE100611) and chromatin-associated RNA-seq data for mock and 8 h p.i. WT infection (GEO accession: GSE100576) were taken from our previous publication (22). H3K36me3 ChIPmentation data for mock and 8 h p.i. WT infection of HFF were also taken from our recent publication (30) (n=3, GEO accession: GSE185246). H2A.Z ChIP-seq data of untreated and α-amanitin-treated HCT116 cells were taken from the study by Lashgari *et al.* (47) (n=2, GSE101427).

ChIPmentation, library preparation and sequencing

HFF were purchased from ECACC and cultured in Dulbecco's Modified Eagle Medium (DMEM, ThermoFisher #41966052) supplemented with 10% (v/v) Fetal Bovine Serum (FBS, Biochrom #S0115), 1× MEM Non-Essential Amino Acids (ThermoFisher #11140050) and 1% penicillin/streptomycin. Two days prior to infection, two million HFF cells were seeded in 15 cm dishes. At the day of infection, cells had expanded to ~80% confluency. Cells were infected with the respective viruses as described in results section (n=3). At 8 p.i., cells were fixed for 10 minutes at room temperature by adding 1% formaldehyde (final) directly to the medium. Cells were scraped in 1mL of ice-cold 1x PBS containing protease inhibitor cocktail (1x) (Roche #11836153001) with an additional 1mM phenylmethylsulfonyl fluoride (PMSF).

Cells were pelleted at 1500 rpm for 20 min at 4 °C. Supernatant was aspirated and cell pellets were frozen in liquid N_2.

Cell pellets were resuspended in 1.5 mL 0.25% [w/v] SDS sonication buffer (10 mM Tris pH=8.0, 0.25% [w/v] SDS, 2 mM EDTA) with 1x protease inhibitors and 1 mM additional PMSF and incubated on ice for 10 minutes. Cells were sonicated in fifteen 1 minute intervals, 25% amplitude, with Branson Ultrasonics SonifierTM S-450 until most fragments were in the range of 200-700 bp as determined by agarose gel electrophoresis. Two million cells used for the preparation of the ChIPmentation libraries were diluted 1:1.5 with equilibration buffer (10 mM Tris, 233 mM NaCl, 1.66% [v/v] Triton X-100, 0.166% [w/v] sodium deoxycholate, 1 mM EDTA, protease inhibitors) and spun at 14,000x g for 10 minutes at 4 °C to pellet insoluble material. Supernatant was transferred to a new 1.5 mL screw-cap tube and topped up with RIPA-LS (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1 mM EDTA pH 8.0, 0.1% [w/v] SDS, 0.1% [w/v] sodium deoxycholate, 1% [v/v] Triton X-100, protease inhibitors) to 200 µL. Input and gel samples were preserved. Lysates were incubated with 1µg/IP of anti-H2A.Z antibody (Diagenode, #C15410201) on a rotator overnight at 4 °C.

Dependent on the added amount of antibody, the amount of Protein A magnetic beads (ThermoFisher Scientific #10001D) was adjusted (e.g. for 1-2 µg of antibody/IP = 15 µL of beads) and blocked overnight with 0.1% [w/v] bovine serum albumin in RIPA buffer. On the following day, beads were added to the IP samples for 2 h on a rotator at 4 °C to capture the antibody-bound fragments. The immunoprecipitated chromatin was subsequently washed twice with 150 µL each of ice-cold buffers RIPA-LS, RIPA-HS (10 mM Tris-HCl pH 8.0, 50 0mM NaCl, 1 mM EDTA pH 8.0, 0.1% [w/v] SDS, 0.1% [v/v] sodium deoxycholate, 1% [v/v] Triton X-100), RIPA-LiCl (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA pH 8.0, 0.5% [w/v] sodium deoxycholate, 0.5% [v/v] Nonidet P-40) and 10 mM Tris pH 8.0

containing protease inhibitors. Beads were washed once more with ice-cold 10 mM Tris pH 8.0 lacking inhibitors and transferred into new tubes.

Beads were resuspended in 25 μL of the tagmentation reaction mix (Nextera DNA Sample Prep Kit, Illumina) containing 5 μL of 5x Tagmentation buffer, 1 μL of Tagment DNA enzyme, topped up with H_2O to the final volume and incubated at 37 °C for 10 minutes in a thermocycler. Beads were mixed after 5 minutes by gentle pipetting. To inactivate the Tn5 enzyme, 150 μL of ice-cold RIPA-LS was added to the tagmentation reaction. Beads were washed twice with 150 μL of RIPA-LS and 1x Tris-EDTA and subjected to de-crosslinking by adding 100 μL ChIPmentation elution buffer (160 mM NaCl, 40 μg/mL Rnase A (Sigma-Aldrich #R4642), 1x Tris-EDTA (Sigma #T9285) and incubating for 1h at 37 °C followed by overnight shaking at 65 °C. The next day, 4 mM EDTA and 200 μg/mL Proteinase K (Roche, #03115828001) were added, and samples incubated for another 2h at 45 °C with 1000 rpm shaking. Supernatant was transferred into a new tube and another 100 μL of ChIPmentation elution buffer was added for another hour at 45 °C with 1000 rpm shaking. DNA was isolated with MinElute PCR Purification Kit (Qiagen #28004) and eluted in 21 μL of H_2O.

DNA for the final library was prepared with 25 μL NEBNext Ultra II Q5 Master Mix, 3.75 μL IDT custom primer i5_n_x (10 μM); 3.75 μL IDT custom primer i7_n_x (10 μM); 3.75 μL H_2O and 13.75 μL ChIPmentation DNA. The Cq value obtained from the library quantification, rounded up to the nearest integer plus one additional cycle, was used to amplify the rest of the ChIPmentation DNA. Library qualities were verified by High Sensitivity DNA Analysis on the Bioanalyzer 2100 (Agilent) before performing sequencing on NextSeq 500 (paired-end 35bp reads) at the Core Unit Systemmedizin, Würzburg, Germany. All samples were sequenced at equimolar ratios.

Read alignment

The read alignment pipeline was implemented and run in the workflow management system Watchdog (57, 58) as already previously described (39). Public sequencing data were downloaded from SRA using the sratoolkit version 2.10.8. Sequencing reads were aligned against the human genome (GRCh37/hg19), the HSV-1 genome (Human herpesvirus 1 strain 17, GenBank accession code: JN555585) and human rRNA sequences using ContextMap2 version 2.7.9 (59) (using BWA as short read aligner (60) and allowing a maximum indel size of 3 and at most 5 mismatches). For the two repeat regions in the HSV-1 genome, only one copy was retained each, excluding nucleotides 1–9,213 and 145,590–152,222 from the alignment. SAM output files of ContextMap2 were converted to BAM files using samtools (61). Read coverage in bedGraph format was calculated from BAM files using BEDTools (62). Subregions of promoter windows (TSS ± 3 kb) with differential read coverage in ATAC-seq and H2A.Z ChIPmentation/-seq data were determined with RegCFinder (37). For this purpose, RegCFinder was applied to ATAC-seq data for all pairwise comparisons of mock to WT and null mutant infections as well as mock to each timepoint of infection for the time-course data and to H2A.Z ChIPmentation and ChIP-seq data for the comparison of mock and WT infection as well as untreated and α-amanitin-treated HCT116 cells.

Data plotting and statistical analysis

All figures were created in R and all statistical analyses were performed in R (63). Read coverage plots were created using the R Bioconductor package Gviz (64).

Metagene and clustering analysis

Metagene analyses were performed as previously described (65) using the R program developed for this previous publication (available with the Watchdog *binGenome* module in the Watchdog module repository (https://github.com/watchdog-wms/watchdog-wms-

modules/)). For promoter region analyses, the regions −3 kb to +3 kb of the TSS were divided into 101 equal-sized bins for each gene. For each bin, the average coverage per genome position was calculated and bin read coverages were then normalized by dividing by the total sum of all bins. Metagene curves for each replicate were created by averaging results for corresponding bins across all genes in a cluster/group and then averaged across replicates. For hierarchical clustering analysis, RegCFinder profiles of differential regions were calculated for each promoter window and comparison by setting each position in an m-RegC to 1, in an i-RegC to -1 and all other positions to 0. RegCFinder profiles of each promoter window for each comparison were concatenated into one row in the matrix. Hierarchical clustering of the resulting matrix was then performed using the hclust function in R according to Euclidean distances and Ward's clustering criterion.

Analysis of downstream open chromatin regions (dOCRs)

Open chromatin regions (OCRs) were determined from ATAC-seq data by first converting BAM files with mapped reads to BED format using BEDTools (66) and then determining enriched regions from these BED files using F-Seq with default parameters (67). dOCRs for individual genes were calculated from OCRs as previously described (22, 30). In brief, dOCRs are determined for genes by first assigning all OCRs overlapping with the 10 kb downstream of a gene to this gene. Second, OCRs starting at most 5 kb downstream of the so far most downstream OCR of a gene are also assigned to this gene. In both steps, individual OCRs can be assigned to multiple genes. The second step is iterated until no more OCRs can be assigned. The dOCR of a gene is then defined as the region from the gene 3'end to the end of the most downstream OCR assigned to this gene.

Gene expression analysis

Number of fragments (=read pairs) per gene were determined from mapped paired-end RNA-seq reads in a strand-specific manner using featureCounts (68) and gene annotations from Ensembl (version 87 for GRCh37). For genes, all fragments overlapping exonic regions on the corresponding strand by ≥ 25bp were counted for the corresponding gene. Fold-changes in gene expression and statistical significance of changes were determined using DESeq2 (69) and p-values were adjusted for multiple testing using the method by Benjamini and Hochberg (70). Gene expression was quantified in terms of fragments per kilobase of exons per million mapped reads (FPKM). Only reads mapped to the human genome were counted for the total number of mapped reads for FPKM calculation.

Motif discovery and enrichment analysis

Motif discovery was performed for each cluster separately for the i- and m-RegC regions using findMotifsGenome.pl script of the Homer suite (40), with our 7,649 input promoter windows as background. For this purpose, we used the hg19 annotation provided by Homer and automated trimming of input windows was disabled. Significant motifs were identified at a q-value (=False Discovery Rate (FDR) calculated with the Benjamini-Hochberg method (70)) cutoff of 0.01. Over-representation of Gene Ontology (GO) terms was performed separately for each cluster using the g:Profiler webserver (71) and the R package gprofiler2 (72), which provides an R interface to the webserver. As background gene list, the genes corresponding to our 7,649 input promoter windows were provided. P-values were corrected for multiple testing using the Benjamini-Hochberg method (70) and significantly over-represented GO terms were identified at an multiple testing adjusted p-value cutoff of 0.001. Enrichment of gene types (obtained from the Ensembl annotation (version 87 for GRCh37)) within clusters was determined using one-sided Fisher's exact tests (with alternative =

greater). P-values were again corrected for multiple testing using the Benjamini-Hochberg method and an adjusted p-value cutoff of 0.001 was used.

References

1. Roizman B KD, RJ W. Herpes Simplex Viruses. In: Knipe DM HP, editor. Fields Virology. 5th. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 2501-601.
2. Pellet PE RB. Herpesviridae. In: Knipe DM HP, editor. Fields Virology. 6th. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013. p. 1802-2128.
3. Ablashi D, Agut H, Alvarez-Lafuente R, Clark DA, Dewhurst S, DiLuca D, et al. Classification of HHV-6A and HHV-6B as distinct viruses. Arch Virol. 2014;159(5):863-70.
4. Korr G, Thamm M, Czogiel I, Poethko-Mueller C, Bremer V, Jansen K. Decreasing seroprevalence of herpes simplex virus type 1 and type 2 in Germany leaves many people susceptible to genital infection: time to raise awareness and enhance control. BMC Infect Dis. 2017;17(1):471.
5. Whitley RJ, Roizman B. Herpes simplex virus infections. Lancet. 2001;357(9267):1513-8.
6. Whisnant AW, Jürges CS, Hennig T, Wyler E, Prusty B, Rutkowski AJ, et al. Integrative functional genomics decodes herpes simplex virus 1. Nat Commun. 2020;11(1):2038.
7. Spencer CA, Dahmus ME, Rice SA. Repression of host RNA polymerase II transcription by herpes simplex virus type 1. J Virol. 1997;71(3):2031-40.
8. Rice SA, Davido DJ. HSV-1 ICP22: hijacking host nuclear functions to enhance viral infection. Future Microbiol. 2013;8(3):311-21.

9. Abrisch RG, Eidem TM, Yakovchuk P, Kugel JF, Goodrich JA. Infection by Herpes Simplex Virus 1 Causes Near-Complete Loss of RNA Polymerase II Occupancy on the Host Cell Genome. J Virol. 2015;90(5):2503-13.

10. Dembowski JA, DeLuca NA. Selective recruitment of nuclear factors to productively replicating herpes simplex virus genomes. PLoS Pathog. 2015;11(5):e1004939.

11. Birkenheuer CH, Danko CG, Baines JD. Herpes Simplex Virus 1 Dramatically Alters Loading and Positioning of RNA Polymerase II on Host Genes Early in Infection. J Virol. 2018;92(8).

12. Rivas HG, Schmaling SK, Gaglia MM. Shutoff of Host Gene Expression in Influenza A Virus and Herpesviruses: Similar Mechanisms and Common Themes. Viruses. 2016;8(4):102.

13. Dremel SE, DeLuca NA. Herpes simplex viral nucleoprotein creates a competitive transcriptional environment facilitating robust viral transcription and host shut off. Elife. 2019;8.

14. McSwiggen DT, Hansen AS, Teves SS, Marie-Nelly H, Hao Y, Heckert AB, et al. Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. eLife. 2019;8:e47098.

15. Dremel SE, DeLuca NA. Herpes simplex viral nucleoprotein creates a competitive transcriptional environment facilitating robust viral transcription and host shut off. eLife. 2019;8:e51109.

16. Isa NF, Bensaude O, Aziz NC, Murphy S. HSV-1 ICP22 Is a Selective Viral Repressor of Cellular RNA Polymerase II-Mediated Transcription Elongation. Vaccines. 2021;9(10):1054.

17. Kwong AD, Frenkel N. Herpes simplex virus-infected cells contain a function(s) that destabilizes both host and viral mRNAs. Proc Natl Acad Sci U S A. 1987;84(7):1926-30.

18. Oroskar AA, Read GS. Control of mRNA stability by the virion host shutoff function of herpes simplex virus. J Virol. 1989;63(5):1897-906.

19. Feng P, Everly DN, Jr., Read GS. mRNA decay during herpesvirus infections: interaction between a putative viral nuclease and a cellular translation factor. J Virol. 2001;75(21):10272-80.

20. Friedel CC, Whisnant AW, Djakovic L, Rutkowski AJ, Friedl MS, Kluge M, et al. Dissecting Herpes Simplex Virus 1-Induced Host Shutoff at the RNA Level. J Virol. 2021;95(3).

21. Rutkowski AJ, Erhard F, L'Hernault A, Bonfert T, Schilhabel M, Crump C, et al. Widespread disruption of host transcription termination in HSV-1 infection. Nat Commun. 2015;6:7126.

22. Hennig T, Michalski M, Rutkowski AJ, Djakovic L, Whisnant AW, Friedl MS, et al. HSV-1-induced disruption of transcription termination resembles a cellular stress response but selectively increases chromatin accessibility downstream of genes. PLoS Pathog. 2018;14(3):e1006954.

23. Wang X, Hennig T, Whisnant AW, Erhard F, Prusty BK, Friedel CC, et al. Herpes simplex virus blocks host transcription termination via the bimodal activities of ICP27. Nat Commun. 2020;11(1):293.

24. Vilborg A, Passarelli MC, Yario TA, Tycowski KT, Steitz JA. Widespread Inducible Transcription Downstream of Human Genes. Mol Cell. 2015;59(3):449-61.

25. Rosa-Mercado NA, Zimmer JT, Apostolidi M, Rinehart J, Simon MD, Steitz JA. Hyperosmotic stress alters the RNA polymerase II interactome and induces readthrough transcription despite widespread transcriptional repression. Mol Cell. 2021;81(3):502-13 e4.

26. Hadar S, Meller A, Saida N, Shalgi R. Stress-induced transcriptional readthrough into neighboring genes is linked to intron retention. iScience. 2022;25(12):105543.

27. Wu N, Watkins SC, Schaffer PA, DeLuca NA. Prolonged gene expression and cell survival after infection by a herpes simplex virus mutant defective in the immediate-early genes encoding ICP4, ICP27, and ICP22. Journal of Virology. 1996;70(9):6358-69.

28. Honess RW, Watson DH. Herpes simplex virus resistance and sensitivity to phosphonoacetic acid. Journal of Virology. 1977;21(2):584-600.

29. Becker Y, Asher Y, Cohen Y, Weinberg-Zahlering E, Shlomai J. Phosphonoacetic Acid-Resistant Mutants of Herpes Simplex Virus: Effect of Phosphonoacetic Acid on Virus Replication and In Vitro Deoxyribonucleic Acid Synthesis in Isolated Nuclei. Antimicrobial Agents and Chemotherapy. 1977;11(5):919-22.

30. Djakovic L, Hennig T, Reinisch K, Milic A, Whisnant AW, Wolf K, et al. The HSV-1 ICP22 protein selectively impairs histone repositioning upon Pol II transcription downstream of genes. Nat Commun. 2023;14(1):4591.

31. Weiner A, Hughes A, Yassour M, Rando OJ, Friedman N. High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. Genome Res. 2010;20(1):90-100.

32. Schwabish MA SK. Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. Mol Cell Biol. 2004;24(23):10111-7.

33. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007;129(4):823-37.

34. Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, et al. Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. Cell. 2005;123(2):233-48.

35. Li S, Wei T, Panchenko AR. Histone variant H2A.Z modulates nucleosome dynamics to promote DNA accessibility. Nature Communications. 2023;14(1):769.

36. Bensaude O. Inhibiting eukaryotic transcription. Which compound to choose? How to evaluate its activity? Transcription. 2011;2(3):103-8.

37. Weiss E, Friedel CC. RegCFinder: targeted discovery of genomic subregions with differential read density. Bioinform Adv. 2023;3(1):vbad085.

38. Anders S RA, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Research. 2012;22(10):2008-17.

39. Weiss E, Hennig T, Grassl P, Djakovic L, Whisnant AW, Jurges CS, et al. HSV-1 Infection Induces a Downstream Shift of Promoter-Proximal Pausing for Host Genes. J Virol. 2023;97(5):e0038123.

40. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.

41. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, et al. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell. 2007;129(4):823-37.

42. Shivaswamy S, Bhinge A, Zhao Y, Jones S, Hirst M, Iyer VR. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. PLoS Biol. 2008;6(3):e65.

43. Tirosh I, Barkai N. Two strategies for gene regulation by promoter nucleosomes. Genome Res. 2008;18(7):1084-91.

44. Hartley PD, Madhani HD. Mechanisms that specify promoter nucleosome location and identity. Cell. 2009;137(3):445-58.

45. Greenberg RS, Long HK, Swigut T, Wysocka J. Single Amino Acid Change Underlies Distinct Roles of H2A.Z Subtypes in Human Syndrome. Cell. 2019;178(6):1421-36 e24.

46. Bagchi DN, Battenhouse AM, Park D, Iyer VR. The histone variant H2A.Z in yeast is almost exclusively incorporated into the +1 nucleosome in the direction of transcription. Nucleic Acids Research. 2020;48(1):157-70.

47. Lashgari A, Millau J-F, Jacques P-É, Gaudreau L. Global inhibition of transcription causes an increase in histone H2A.Z incorporation within gene bodies. Nucleic Acids Research. 2017;45(22):12715-22.

48. Wyler E, Menegatti J, Franke V, Kocks C, Boltengagen A, Hennig T, et al. Widespread activation of antisense transcription of the host genome during herpes simplex virus 1 infection. Genome Biology. 2017;18(1):209.

49. Jimeno-González S, Ceballos-Chávez M, Reyes JC. A positioned +1 nucleosome enhances promoter-proximal pausing. Nucleic Acids Research. 2015;43(6):3068-78.

50. Conn KL HM, Schang LM. Linker histones are mobilized during infection with herpes simplex virus type 1. J Virol. 2008;82:8629–46.

51. Conn KL HM, Schang LM. Core histones H2B and H4 are mobilized during infection with herpes simplex virus 1. J Virol. 2011;85:13234–52.

52. Conn KL, Schang LM. Chromatin Dynamics during Lytic Infection with Herpes Simplex Virus 1. Viruses. 2013;5(7):1758-86.

53. Stow ND, Stow EC. Isolation and Characterization of a Herpes Simplex Virus Type 1 Mutant Containing a Deletion within the Gene Encoding the Immediate Early Polypeptide Vmw110. Journal of General Virology. 1986;67(12):2571-85.

54. Post LE, Roizman B. A generalized technique for deletion of specific genes in large genomes: a gene 22 of herpes simplex virus 1 is not essential for growth. Cell. 1981;25(1):227-32.

55. Smith IL, Hardwicke MA, Sandri-Goldin RM. Evidence that the herpes simplex virus immediate early protein ICP27 acts post-transcriptionally during infection to regulate gene expression. Virology. 1992;186(1):74-86.

56. Fenwick ML, Everett RD. Inactivation of the Shutoff Gene (UL41) of Herpes Simplex Virus Types 1 and 2. Journal of General Virology. 1990;71(12):2961-7.

57. Kluge M, Friedel CC. Watchdog - a workflow management system for the distributed analysis of large-scale experimental data. BMC Bioinformatics. 2018;19(1):97.

58. Kluge M, Friedl MS, Menzel AL, Friedel CC. Watchdog 2.0: New developments for reusability, reproducibility, and workflow execution. Gigascience. 2020;9(6).

59. Bonfert T, Kirner E, Csaba G, Zimmer R, Friedel CC. ContextMap 2: fast and accurate context-based RNA-seq mapping. BMC Bioinformatics. 2015;16:122.

60. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.

61. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2).

62. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.

63. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria2022.

64. Hahne F, Ivanek R. Visualizing Genomic Data Using Gviz and Bioconductor. Methods Mol Biol. 2016;1418:335-51.

65. Chirackal Manavalan AP, Pilarova K, Kluge M, Bartholomeeusen K, Rajecky M, Oppelt J, et al. CDK12 controls G1/S progression by regulating RNAPII processivity at core DNA replication genes. EMBO Rep. 2019;20(9):e47592.

66. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.

67. Boyle AP, Guinney J, Crawford GE, Furey TS. F-Seq: a feature density estimator for high-throughput sequence tags. Bioinformatics. 2008;24(21):2537-8.

68. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.

69. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

70. Benjamini Y HY. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.

71. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47(W1):W191-W8.

72. Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. F1000Res. 2020;9.

Supporting information

S1 File. Supplementary Figures.

Contains Supplementary Figures 1-26 and legends (PDF).

Figures

Fig 1: HSV-1 infection impacts chromatin accessibility around promoters for the majority of host genes.

(a) Heatmap illustrating the location of differential regions (m- and i-RegC) identified by RegCFinder on the ATAC-seq data for WT, ΔICP0, ΔICP22, ΔICP27, and Δ*vhs* infection in comparison to mock infection. Results are shown for 4,981 (out of 7,649) promoter windows containing at least one statistically significant (adj. p. ≤ 0.01) differential region for at least one virus infection compared to mock. Red and blue colors represent m-RegC (mock>infection) and i-RegC (infection>mock) locations, respectively, within promoter windows. Here, each row of the heatmap shows results for the same input window in different comparisons of virus infection to mock. Colored rectangles on top of the heatmap indicate which virus infection was compared to mock. Black vertical lines in the center of each part of the heatmap indicate the position of the TSS. Hierarchical clustering was performed according to Euclidean distances and Ward's clustering criterion and the cutoff on the hierarchical clustering dendrogram was selected to obtain 14 clusters (marked by colored rectangles between dendrogram and heatmap and numbered from top to bottom as indicated). **(b)** Metagene plot showing the average ATAC-seq profile around the TSS ± 3 kb in mock (red) and WT (blue) infection for cluster 5, an example for pattern (i). For a description of metagene plots, see Materials and Methods. The colored bands below the metagene curves in each panel indicate the percentage of genes having an m- or i-RegC (red or blue, respectively)

at that position. **(c)** Read coverage in a ±3.6 kb window around the TSS in ATAC-seq data of mock (red) and WT (blue) infection for an example gene (FBXO28) with pattern (i). Read coverage was normalized to total number of mapped reads for each sample and averaged between replicates. The TSS is indicated by a short vertical line below the read coverage track for mock infection. Gene annotation is indicated at the top. Boxes represent exons, lines represent introns, and direction of transcription is indicated by arrowheads. Below the read coverage track for WT infection, m-RegC (red bars) and i-RegC (blue bars) are shown for the comparison of WT vs. mock infection. **(d)** Metagene plot as in **(b)** for cluster 7, the most pronounced case of pattern (ii). **(e)** Read coverage plot as in **(c)** for an example gene (ARHGAP1) with pattern (ii). Here, the promoter window also contains the TSS of the ZNF408 gene on the opposite strand and the accessible chromatin region extends upstream of the ARHGAP1 TSS, i.e. downstream of the ZNF408 TSS. **(f,g)** Metagene plots as in **(b)** for **(f)** cluster 13, which exhibits pattern (iii), and **(g)** cluster 8, one of two clusters exhibiting the combined (i)+(ii) pattern.

Fig 2: Changes in chromatin accessibility at promoters are mostly independent of dOCR induction.

(a) Bar plots showing the percentage of promoter regions for each cluster overlapping with dOCRs downstream of the 1,296 genes for which we previously showed consistent dOCR induction across different HSV-1 strains (30). dOCR regions in mock, WT (± PAA treatment) and null mutant infection were calculated as previous described (30) (see Materials and Methods for further details) and the overlaps of promoter windows to dOCR regions originating from an upstream gene were determined for each cluster. WT (± PAA treatment) and null mutant infection are ordered according to the overall extent of dOCR induction. Only few clusters (12-14) showed some enrichment for dOCRs from upstream genes, which was

abolished in ΔICP22 infection. **(b-c)** Metagene curves showing average ATAC-seq profiles in promoter windows for mock (red), WT (blue), WT+PAA (green) and ΔICP22 (orange) infection for clusters 5 **(b)** and 7 **(c)**. See Materials and Methods for a detailed description of metagene plots. (**d**) Barplot showing the percentage of genes in each cluster annotated as either protein-coding, long intervening/intergenic noncoding RNAs (lincRNA), snRNA, antisense, or other in Ensembl. Significance of enrichment of each gene type in each cluster was determined using a one-sided Fisher's exact test (with alternative = greater) and p-values were corrected for multiple testing using the Benjamini-Hochberg method. Adj. p. < 0.05 are indicated in the corresponding field of the barplot. (**e**) Percentage of bidirectional promoters in each cluster was calculated as the percentage of promoter windows containing a protein-coding, lincRNA or antisense gene (according to Ensembl annotation) on the opposite strand to the target gene starting within 1 kb upstream of the TSS of the target gene. Significance analysis and multiple testing correction was performed as for (**d**) and adj. p. < 0.05 are indicated on top of bars.

Fig 3: Changes in chromatin accessibility at promoters are linked to transcription.

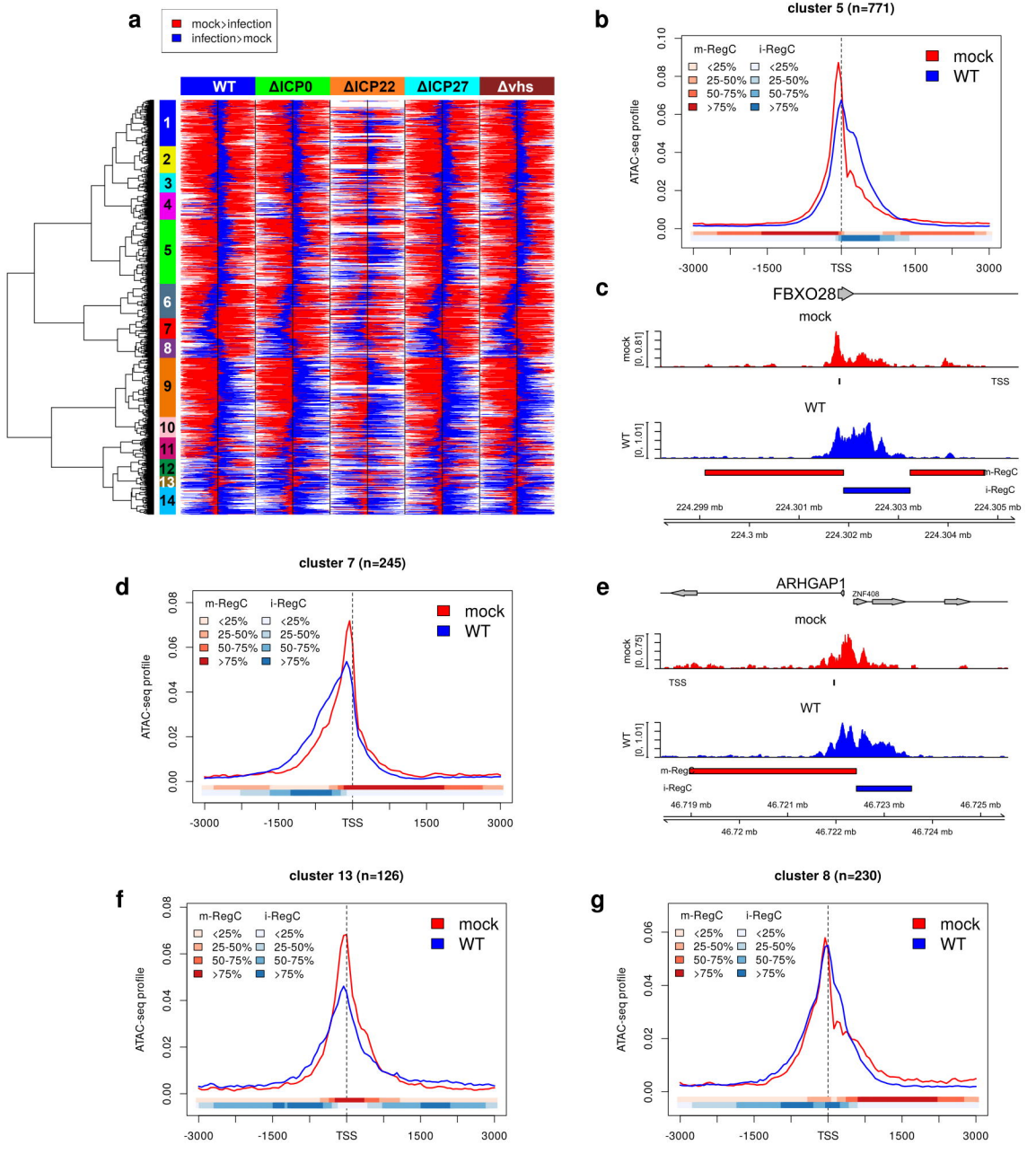
(a,b) Boxplots showing the distribution of **(a)** log2(DSR:UAR) ratios and **(b)** gene expression (FPKM) values in chromatin-associated RNA in mock (red) and 8 h p.i. WT (blue) HSV-1 infection for all clusters (grouped by pattern, cluster numbers are shown below each boxplot) and remaining genes without significant chromatin accessibility changes (NA group). DSR was calculated as the expression in chromatin-associated RNA for the region downstream of the TSS in sense direction of the target gene and UAR as the expression in the upstream region in antisense direction. P-values for Wilcoxon rank sum tests comparing values in mock infection for each cluster/group against all other analyzed genes are indicated below cluster numbers and were corrected for multiple testing using the Bonferroni method. **(c)** Heatmap showing the log2 fold-changes determined with DEXSeq on the ATAC-seq time-course data

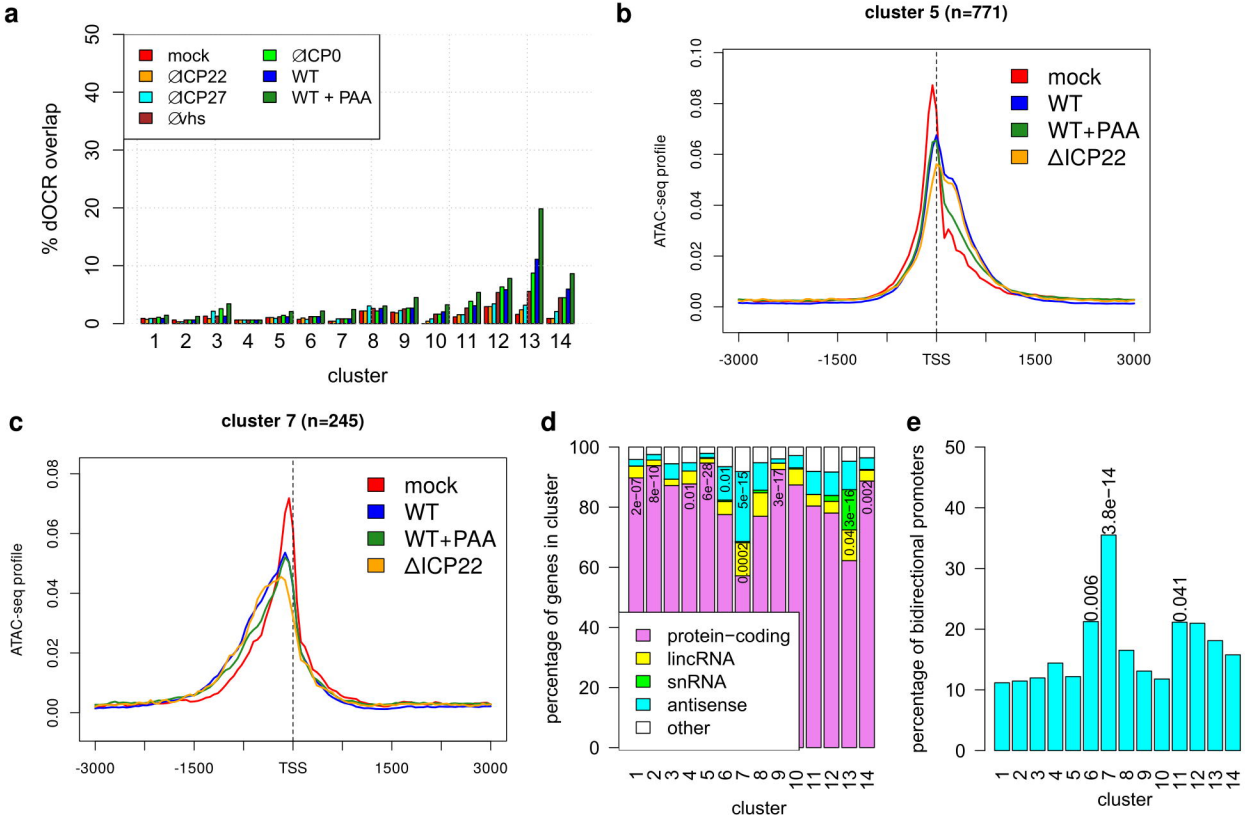
(1, 2, 4, 6 and 8 h p.i. WT infection compared to mock infection) for statistically significant differential regions (m- and i-RegC) identified by RegCFinder. For comparison, log2 fold-changes for the WT vs. mock comparison from **Fig 1a** are also shown. Each row represents the results for one of the input windows included in **Fig 1a**. Colored rectangles on top indicate the time-point of infection or whether the original WT vs. mock comparison is shown. Statistically significant differential regions are colored according to the log2 fold-change determined by DEXSeq. Here, the color scale is continuous between -1 and 1 and all log2 fold-changes >1 are colored the same red and all log2 fold-changes <1 the same blue. Promoter windows are ordered as in **Fig 1a** and clusters are annotated as colored and numbered rectangles on the left. **(d)** Number of statistically significant differential regions (m- and i-RegC) and number of genes with at least one statistically significant differential region identified for each time-point of infection in the ATAC-seq time-course data. **(e)** Metagene plot of ATAC-seq profiles in mock infection (red) and all time-points of infection (blue shades) from the ATAC-seq time-course experiment for example cluster 5 (pattern (i)). The colored bands below the metagene curves indicate the percentage of genes having an i- or m-RegC (blue or red, respectively) at that position in the comparison of 8 h p.i. to mock infection from the time-course experiment. Metagene plots for all clusters are shown in **Sup. Fig 20** in **S1 File**.

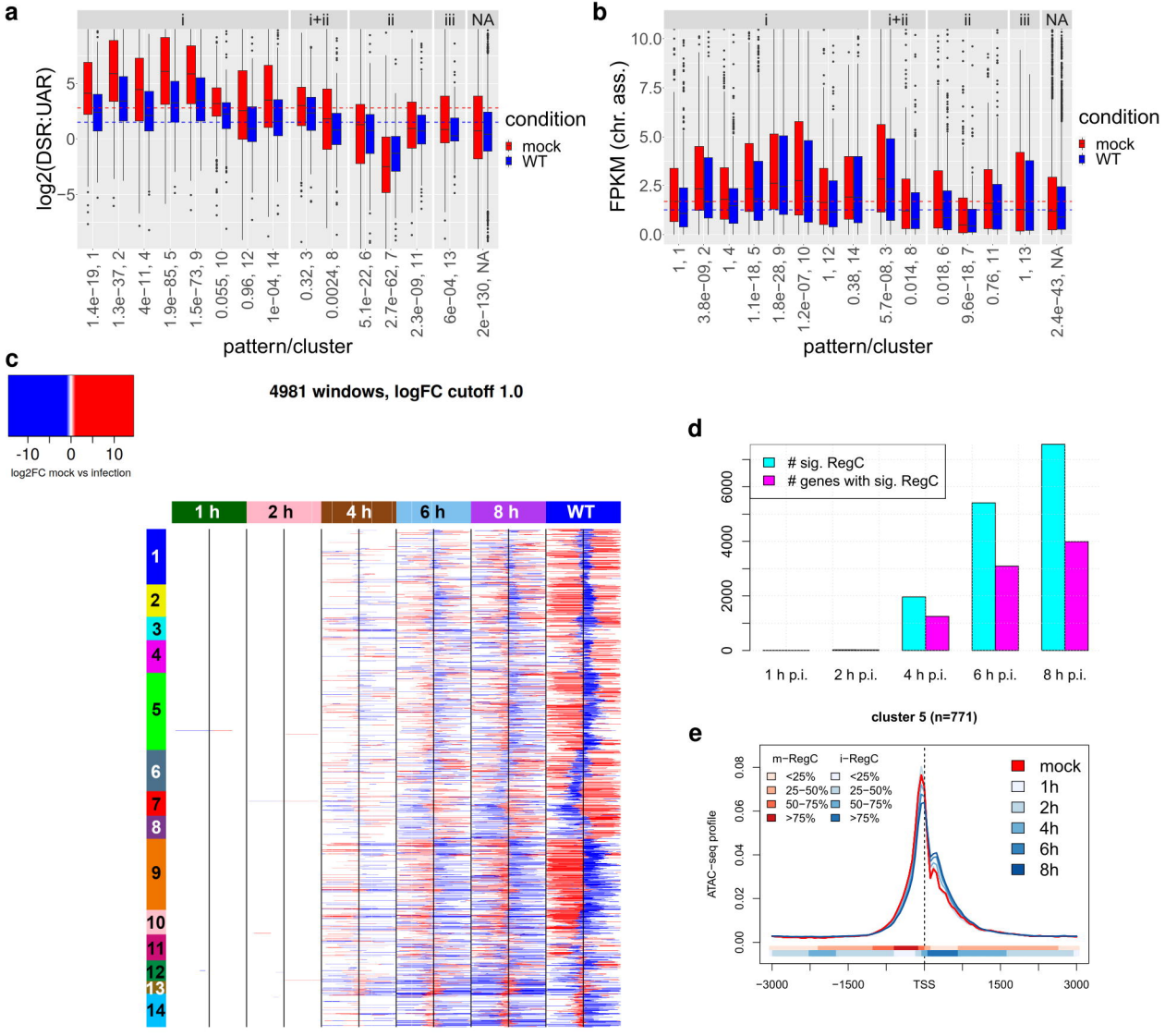
Fig 4: HSV-1 infection and depletion of Pol II from human genes lead to a downstream shift of H2A.Z-containing +1 nucleosomes.

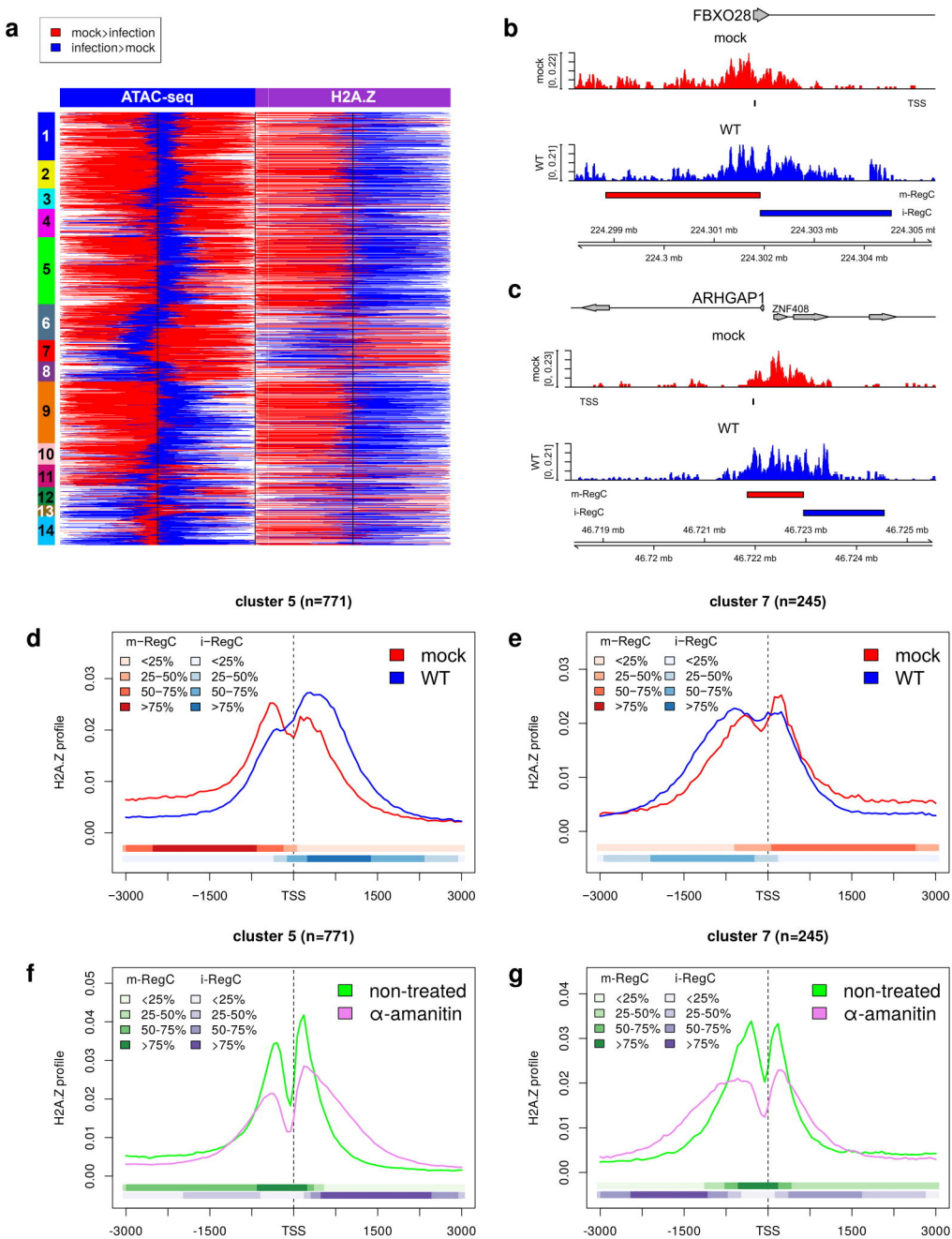
(a) Heatmap showing the location for differential regions (m- and i-RegC) identified in the WT vs. mock comparison on the ATAC-seq (left half, identical to left-most part of **Fig 1a**) and H2A.Z ChIPmentation data (right half) for the promoter windows included in **Fig 1a.** Differential regions with mock>infection (m-RegC) are marked in red, differential regions

with infection>mock (i-RegC) in blue. The order of promoter windows is the same as in **Fig 1a** and clusters are indicated by colored and numbered rectangles on the left. **(b,c)** Read coverage in a ±3.6 kb window around the TSS in H2A.Z ChIPmentation data for mock (red) and WT (blue) infection for example genes with **(b)** pattern (i) (FBXO28) and **(c)** pattern (ii) (ARHGAP1). For descriptions of read coverage plots see caption to **Fig 1**. **(d-g)** Metagene plots of H2A.Z profiles for mock and WT infection **(d,e)** and untreated and α-amanitin-treated HCT116 cells **(f,g)** for clusters 5 (pattern (i), **d,f**) and 7 (pattern (ii), **e,g**). Metagene plots for other clusters can be found in **Sup. Fig 23** and **25** in **S1 File**, respectively. The colored bands below the metagene curves in each panel indicate the percentage of genes having m- or i-RegC at that position. For **(f,g)**, m-RegC (green) are differential regions with relative read coverage higher in untreated cells and i-RegC (violet) differential regions with relative read coverage higher upon α-amanitin-treatment.


a
mock>infection
infection>mock
WT
ΔICP0
ΔICP22
ΔICP27
Δvhs
1
2
3
4
5
6
7
8
9
10
11
12
13
14
b
cluster 5 (n=771)
m-RegC
i-RegC
<25%
25-50%
50-75%
>75%
mock
WT
ATAC-seq profile
-3000
-1500
TSS
1500
3000
c
FBXO28
mock
WT
TSS
m-RegC
i-RegC
224.299 mb
224.301 mb
224.303 mb
224.305 mb
224.3 mb
224.302 mb
224.304 mb
d
cluster 7 (n=245)
e
ARHGAP1
ZNF408
46.719 mb
46.721 mb
46.723 mb
46.725 mb
46.72 mb
46.722 mb
46.724 mb
f
cluster 13 (n=126)
g
cluster 8 (n=230)


a
mock
∅ICP22
∅ICP27
∅vhs
∅ICP0
WT
WT + PAA
% dOCR overlap
cluster
b
cluster 5 (n=771)
mock
WT
WT+PAA
ΔICP22
ATAC-seq profile
-3000
-1500
TSS
1500
3000
c
cluster 7 (n=245)
mock
WT
WT+PAA
ΔICP22
ATAC-seq profile
d
percentage of genes in cluster
protein−coding
lincRNA
snRNA
antisense
other
cluster
e
percentage of bidirectional promoters
0.006
3.8e−14
0.041
cluster


a
i
i+ii
ii
iii
NA
log2(DSR:UAR)
condition
mock
WT
1.4e−19, 1
1.3e−37, 2
4e−11, 4
1.9e−85, 5
1.5e−73, 9
0.055, 10
0.96, 12
1e−04, 14
0.32, 3
0.0024, 8
5.1e−22, 6
2.7e−62, 7
2.3e−09, 11
6e−04, 13
2e−130, NA
pattern/cluster
b
FPKM (chr. ass.)
1, 1
3.8e−09, 2
1, 4
1.1e−18, 5
1.8e−28, 9
1.2e−07, 10
1, 12
0.38, 14
5.7e−08, 3
0.014, 8
0.018, 6
9.6e−18, 7
0.76, 11
1, 13
2.4e−43, NA
c
4981 windows, logFC cutoff 1.0
log2FC mock vs infection
1 h
2 h
4 h
6 h
8 h
WT
d
sig. RegC
genes with sig. RegC
1 h p.i.
2 h p.i.
4 h p.i.
6 h p.i.
8 h p.i.
e
cluster 5 (n=771)
m−RegC
i−RegC
<25%
25−50%
50−75%
>75%
mock
1h
2h
4h
6h
8h
ATAC−seq profile
TSS


a
mock>infection
infection>mock
ATAC-seq
H2A.Z
1
2
3
4
5
6
7
8
9
10
11
12
13
14
b
FBXO28
mock
mock
[0, 0.22]
TSS
WT
WT
[0, 0.21]
m-RegC
i-RegC
224.299 mb
224.301 mb
224.303 mb
224.305 mb
224.3 mb
224.302 mb
224.304 mb
c
ARHGAP1
ZNF408
mock
mock
[0, 0.23]
TSS
WT
WT
[0, 0.21]
m-RegC
i-RegC
46.719 mb
46.721 mb
46.723 mb
46.725 mb
46.72 mb
46.722 mb
46.724 mb
cluster 5 (n=771)
cluster 7 (n=245)
d
e
m-RegC
i-RegC
<25%
25−50%
50−75%
>75%
mock
WT
H2A.Z profile
0.00
0.01
0.02
0.03
−3000
−1500
TSS
1500
3000
cluster 5 (n=771)
cluster 7 (n=245)
f
g
m-RegC
i-RegC
<25%
25-50%
50-75%
>75%
non-treated
α-amanitin
H2A.Z profile
0.00
0.01
0.02
0.03
0.04
0.05
-3000
-1500
TSS
1500
3000