HSV-1 infection induces a downstream shift of promoter-proximal pausing linked to H2A.Z-containing +1 nucleosomes

Elena Weiß[1], Thomas Hennig[2], Pilar Grassl[1], Lara Djakovic[2], Adam W. Whisnant[2], Christopher S. Jürges[2], Franziska Koller[1], Michael Kluge[1], Florian Erhard[2], Lars Dölken[2,3#], 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] Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz-Center for Infection Research (HZI), 97080 Würzburg, Germany

[#] Corresponding authors

Abstract

Herpes simplex virus 1 (HSV-1) infection exerts a profound shut-off of host gene expression at multiple levels. Recently, HSV-1 infection was reported to also impact promoter-proximal RNA Polymerase II (Pol II) pausing, a key step in the eukaryotic transcription cycle. Here, we reveal that HSV-1 infection induces complex alterations in promoter-proximal pausing for the vast majority of cellular genes. While pausing is retained for most host genes in HSV-1 infection, it is shifted to more downstream and less well-positioned sites. This is associated with an extension of the nucleosome-free region at promoters into downstream regions, which did not require HSV-1 late gene expression, the immediate-early HSV-1 proteins ICP0, ICP4, ICP22 and ICP27 or the virion-host shutoff (*vhs*) protein. The downstream shift in pausing coincides with a downstream shift of the noncanonical histone variant H2A.Z associated with +1 nucleosomes. Our data suggest a model in which HSV-1 infection induces a loss of promoter-proximal pausing at the first NELF-dependent pause site. This delays pausing to the second +1 nucleosome-associated pause site, which is shifted downstream in HSV-1 infection potentially due to the loss of Pol II transcription on the cellular genome in HSV-1 infection.

Introduction

Lytic Herpes simplex virus 1 (HSV-1) infection exerts a profound shutoff of host gene expression via degradation of host and viral mRNAs mediated by the HSV-1 virus host shutoff protein (*vhs)* (1, 2) and inhibition of host transcriptional activity (3-6). Efficient recruitment of RNA polymerase II (Pol II) and elongation factors from the host chromatin to replicating viral genomes leads to a substantial loss of Pol II occupancy from the host genome as early as 2-3 h post infection (h p.i.) (3-6). By 8 h p.i., host transcriptional activity is estimated to be only 10-20% of uninfected cells (7). We previously showed that HSV-1 infection disrupts transcription termination for the majority but not all cellular genes, leading to read-through transcription for tens-of-thousands of nucleotides beyond the poly(A) site (8). More recently, Rivas *et al.* revealed that HSV-1 also impacts promoter-proximal pausing of Pol II on host genes (9), a key step in the eukaryotic transcription cycle. Following transcription initiation, Pol II pauses 20-60 nucleotides downstream of the transcription start site (TSS) (10, 11) as a consequence of one or more structural rearrangements within the transcription elongation complex (12). Pausing makes Pol II vulnerable to nucleosome-induced arrest, backtracking of the elongation complex along the DNA, and promoter-proximal premature termination (12). The elongation factor TFIIS can rescue Pol II from pausing and restart transcription by mediating cleavage of backtracked RNA (13). In contrast, 5,6-dichloro-1-beta-d-ribofuranosylbenzimidazole sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) stabilize paused Pol II (14, 15). Phosphorylation of DSIF, NELF and the Pol II C-terminal domain by the CDK9 subunit of the positive transcription elongation factor b (P-TEFb) is required for the release of paused Pol II into gene bodies and the switch to productive elongation (16-18). As a consequence, inhibition of P-TEFb increases promoter-proximal pausing (19). The HSV-1 immediate-early protein ICP22 inhibits Pol II transcription elongation by direct interaction with CDK9 (20) and ectopic expression of a short segment of ICP22 mimics the effects of P-TEFb inhibition on Pol II transcription (21). However, Rivas *et al.* reported based on Pol II ChIP-seq that HSV-1 infection actually decreased promoter-proximal Pol II pausing for activated genes, which was largely dependent on ICP4 (9). ICP4 is one of five immediate-early proteins (including also ICP0, ICP22, ICP27, and ICP47) expressed shortly after infection and is necessary for transcription of early and late viral genes (22). HSV-1-activated genes exhibited a greater increase in Pol II occupancy on gene bodies than on promoters, consistent with increased transcriptional elongation. For repressed genes, promoter-proximal Pol II pausing was

increased independently of ICP4 with Pol II occupancy decreasing more strongly on gene bodies than in the promoter region.

Here, we report on a genome-wide investigation on the impact of HSV-1 infection on promoter-proximal pausing of all expressed host genes. This is based on a re-analysis of precision nuclear run-on analysis (PRO-seq) data for mock and 3 h p.i. HSV-1 strain F (WT-F) infection from a previous study by Birkenheuer *et al.* (5). PRO-seq sequences RNA that is actively transcribed by Pol II and depicts strand-specific Pol II transcriptional activity. Transcription initiation from most human gene promoters is bidirectional with productive transcription elongation occurring only in sense direction (23-26). PRO-seq thus provides nucleotide-level resolution of Pol II activity and allows separating sense and antisense Pol II initiation and pausing. Birkenheuer *et al.* already reported that Pol II levels at the promoter-proximal pause site were altered in a gene-specific manner (5), but did not explicitly investigate changes in Pol II pausing. Our re-analysis uncovered that HSV-1 infection does not simply lead to a loss or reduction of Pol II pausing for the majority of genes but rather shifts Pol II pausing to sites further downstream from the promoter. In contrast to well-defined Pol II pausing peaks observed in mock infection at the TSS, HSV-1 infection resulted in more varied and less well-positioned patterns of Pol II pausing. This included both broadening of Pol II pausing peaks into downstream regions for some genes as well as newly originating or increasing Pol II peaks at downstream sites for other genes.

Previous studies indicated that Pol II pausing at promoters supports maintenance of the nucleosome-free region at promoters keeping them accessible to transcription factors (27, 28). Using analysis of chromatin accessibility using ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing), we reveal that the downstream shift in promoter-proximal Pol II pausing in HSV-1 infection coincided with increased chromatin accessibility downstream of the TSS that extended the nucleosome-free region at the promoter. This was also observed in absence of ICP4, ICP0, ICP22, ICP27, or the HSV-1 virion-host shutoff (*vhs*) protein as well as upon inhibition of the viral DNA polymerase by phosphonoacetic acid (PAA), which prevents HSV-1 late gene expression. Furthermore, it was accompanied by a downstream shift in occupancy of the noncanonical histone H2A variant H2A.Z associated with +1 nucleosomes. H2A.Z is highly enriched at gene promoters at -1 and +1 nucleosome positions (29, 30) and its deposition at promoters is positively correlated to Pol II pausing (31). H2A.Z has previously been associated with maintenance of Pol II pausing (32) and has been reported

to be mobilized in HSV-1-infected cells (33, 34). Our data thus link the downstream shift of promoter-proximal Pol II pausing sites upon HSV-1 infection to downstream shifts of the H2A.Z-containing +1 nucleosomes.

Results

Widespread changes in promoter-proximal Pol II pausing during HSV-1 infection

The standard measure for quantifying promoter-proximal pausing is the so-called pausing index (PI) of a gene, which is calculated as the ratio of normalized read counts in a window around the TSS (=promoter window) divided by normalized read counts in a window on the gene body excluding the promoter. PIs were also used by Rivas *et al*. to quantify the effects of lytic HSV-1 infection on Pol II pausing (9). We thus started by performing a genome-wide PI analysis using the published PRO-seq data of mock and 3 h p.i. WT-F infection (5). Since annotated gene 5'ends do not necessarily reflect the used TSS in a cell type and multiple alternative TSS are often annotated, we first identified the dominantly used TSS for each gene from published PROcap-seq and PRO-seq data of flavopiridol-treated uninfected human foreskin fibroblasts (HFF) (35) (see methods). PROcap-seq is a variation of PRO-seq that specifically maps Pol II initiation sites. Flavopiridol inhibits CDK9 and thus arrests Pol II in a paused state at the TSS (30) and allows also identifying the TSS for genes that are not or weakly paused in untreated cells. Consistent peaks in PROcap-seq and PRO-seq of flavopiridol-treated HFF provided an initial set of 136,090 putative TSS positions, which were further filtered to identify high confidence sites by requiring a maximum distance of 500 bp to the nearest annotated gene. This identified 42,193 potential TSS positions for 7,650 genes (median number of TSS per gene = 4 with a median distance of 42 bp). For each gene, the TSS with the highest expression was selected for further analyses. Although the PRO-seq data by Birkenheuer *et al.* was obtained in HEp-2 cells, for most genes the identified TSS in HFF matched very well to PRO-seq peaks in mock infected HEp-2 cells (**Sup. Fig. 1a**), better than annotated gene 3'ends (**Sup. Fig. 1b**).

For PI calculation, normalized read counts were determined in a strand-specific manner as reads per kilobase million (RPKM) in the window from the TSS to TSS + 250 bp for the promoter region and from TSS + 250 bp to TSS +2,250 bp (or the gene 3'end if closer) for the gene body. It should be noted that there is no consensus on how to best define promoter and gene body windows for PI calculation and a wide range of alternative ranges have previously been used (see e.g., (31, 36, 37)). Genes with zero reads in the promoter or gene body window

in mock or WT-F 3 h p.i. were excluded, resulting in PIs for 7,056 genes (**Sup. Table 1**). Consistent with previous reports (38), PI was weakly but significantly negatively correlated to gene expression (i.e., gene body RPKM) for both mock and HSV-1 infection (spearman rank correlation $R < -0.25$, p-value for $R \neq 0 < 10^{-98}$, **Sup. Fig. 2a,b**). Comparison of PIs between mock and HSV-1 infection showed a PI reduction for most genes during HSV-1 infection (**Fig. 1a**). An increased PI in HSV-1 infection (fold-change >1 in HSV-1 infection compared to mock) was observed only for 763 genes (red in **Fig. 1a**), a slightly reduced PI (fold-change < 1 but ≥ 0.5) for 2,082 genes (blue in **Fig. 1a**), but a strongly reduced PI (fold-change < 0.5) was observed for 4,211 genes (green in **Fig. 1a**). Genes with a strong PI reduction in HSV-1 infection exhibited the highest PIs in uninfected cells, while genes with a slight PI increase in HSV-1 infection tended to have lowest PIs in uninfected cells (**Fig. 1b**). This trend was reversed in HSV-1 infection with lowest PIs observed for genes with the highest PI reduction.

Genes with either a PI increase or a strong PI decrease upon HSV-1 infection were similarly expressed in uninfected cells, but showed significantly lower expression than genes with a slight PI reduction (Wilcoxon rank sum test $p < 10^{-5}$, **Fig. 1c**). In contrast, genes with a strong PI reduction were more highly expressed in HSV-1 infection than the two other groups and gene expression in HSV-1 infection was weakly negatively correlated to PI fold-changes ($R = -0.18$, p-value for $R \neq 0 < 10^{-52}$, **Sup. Fig. 2c**). This is consistent with the observation by Rivas *et al.* that PIs were reduced for activated genes (9) and suggests that reduction of Pol II pausing contributes to up-regulation in HSV-1 infection or at least partly compensates for the loss of host transcriptional activity. We therefore investigated changes in *de novo* transcription during the first 8 h of HSV-1 infection using our published 4sU-seq time-course for every hour of the first 8 h of lytic HSV-1 strain 17 (WT-17) infection of human fetal foreskin fibroblasts (HFFF) (8). 4sU-seq sequences newly transcribed RNA obtained by labeling with 4-thiouridine (4sU) in specific time intervals of infection (here: 1 h intervals). Recently, we showed that total RNA fold-changes are strongly biased by global changes in RNA stability and transcription during HSV-1 infection (7). Fold-changes in 4sU-RNA, in contrast, provide a largely unbiased picture of gene expression changes. Fold-changes in 4sU-RNA were calculated using DESeq2. In the context of a global reduction in gene expression as observed for HSV-1 infection, these fold-changes represent only relative changes between genes. Accordingly, positive log2 fold-changes do not necessarily imply an increase in transcription but could also represent less transcriptional loss than for other genes. We previously showed that late in infection so-called “read-in” transcription originating from

disrupted transcription termination for an upstream gene that extends into a downstream gene can be mistaken for induction of the downstream gene (8). However, read-in transcription is largely absent within the first 4 h p.i., thus we did not further filter genes based on presence of read-in transcription at later times. From 1-2 h p.i. onwards, genes with a strong PI reduction in HSV-1 infection showed highest median log2 fold-changes, i.e. an increase or less reduction in transcription, while genes with a PI increase exhibited lowest median log2 fold-changes, i.e. the most reduction in transcription (**Fig. 1d**). Our results are thus consistent with the findings by Rivas *et al*. that genes activated during HSV-1 infection showed reduced PI compared to mock infection (9).

Although the HSV-1 protein ICP22 interacts with CDK9 (20), a key regulator of promoter-proximal pausing (16-18), ectopic expression of an ICP22 segment mimics the effects of CDK9 inhibition (21), which arrests Pol II at the promoter rather than reduces pausing (19). Thus, the ICP22-CDK9 interaction does not explain the observed widespread reduction in PI during HSV-1 infection. However, inhibition of a different CDK, i.e., CDK7, has recently been shown to lead to a reduction in Pol II pausing using PRO-seq of 15 min, 30 min, 1 h and 2 h treatment with THZ1, a highly selective covalent CDK7 inhibitor, compared to 1 h DMSO treatment as control (39). As LDC4297A, a different selective inhibitor of CDK7, has been shown to have strong antiviral activity against HSV-1 (40) and CDK7 is also a target of CDK inhibitors inhibiting HSV-1 replication (41), CDK7 likely plays a role of in HSV-1 infection. To investigate whether PI reductions in HSV-1 infection were similar to PI reductions upon CDK7 inhibition, we calculated PIs from the PRO-seq data of CDK7 inhibition for the same genes and TSS as analyzed in our study. PIs correlated moderately but highly significantly between mock-infected and DMSO-treated cells (**Sup. Fig. 3a**, Spearman correlation R=0.50, p-value for $R \neq 0 < 10^{-15}$). However, reductions in PIs were less pronounced upon CDK7 inhibition (**Sup. Fig. 3b**) and fewer genes showed a reduced PI upon CDK7 inhibition than upon HSV-1 infection (**Sup. Fig. 3c**). Moreover, fold-changes in PI after 1 h and 2 h CDK7 inhibition were only weakly (but highly significantly) correlated to fold-changes after HSV-1 infection (**Sup. Fig. 4**, R=0.22 for 2 h CDK7 inhibition, p-value for $R \neq 0 < 10^{-15}$). In summary, this demonstrates that HSV-1 infection impacts promoter-proximal Pol II pausing more profoundly than ≤ 2 h CDK7 inhibition, with small similarities between affected genes.

HSV-1 infection shifts Pol II pausing to downstream sites

A disadvantage of PI analysis is that PIs are not only impacted by increases or decreases in Pol II pausing but by any alteration in Pol II occupancy affecting the number of reads in either promoter or gene body windows. We thus next investigated read coverage in a genome viewer for several example genes with reduced PI. This indicated that changes in promoter-proximal Pol II pausing during HSV-1 infection were highly complex as exemplified by the ATP5G1 and METTL13 genes in **Fig. 2a**. Instead of narrow promoter-proximal PRO-seq peaks observed in mock infection, promoter peaks in HSV-1 infection often extended into the gene body by a few hundred nucleotides (e.g., ATP5G1) and/or additional downstream peaks were observed as e.g., for METTL13. However, at the end of these extended/additional peaks read levels dropped again to similarly low levels (relative to the promoter peak) as in uninfected cells. Read levels were not increased across the whole gene body as would be expected with increased levels of elongation and productive transcription. Since the extended/additional peaks in HSV-1 infection commonly extended >250 bp from the TSS, gene body RPKM, i.e. the denominator in PI calculation, was increased and thus PI values were reduced compared to mock infection.

To investigate whether this was a global trend, we performed metagene analyses in ± 3 kb windows around promoters for all 7,650 analyzed genes (**Sup. Fig. 5a**) as well as separately for the three gene groups defined above based on PI changes (**Fig. 2b-d**). For metagene analyses, the 6 kb promoter windows for each gene were divided into 101 bins, PRO-seq read counts were determined for each bin in a strand-specific manner, normalized to sequencing depth and averaged across replicates. Subsequently, bin values for each gene were normalized to sum up to 1 to obtain the Pol II occupancy profile around the promoter for each gene before averaging across all genes. This normalization allows comparing Pol II occupancy around the promoter between genes with different expression levels and makes the analysis independent of global changes in Pol II occupancy between mock and HSV-1 infection. Please note that normalization was performed independently for sense and antisense transcription, thus the height of peaks is not comparable between sense of antisense transcription. Genes without reads in any samples analyzed were omitted from metagene analyses (=1-2 genes, depending on conditions). Analysis of all genes already showed that PRO-seq profiles for both sense and antisense direction were significantly altered between mock and HSV-1 infection (**Sup. Fig. 5a**). In HSV-1 infection, lower Pol II occupancy was observed directly at the TSS and increased occupancy down- and upstream of the TSS for sense and antisense transcription, respectively. This was limited to within 2,500 bp of the TSS in both cases. Metagene analysis

on complete genes from the promoter to downstream of the transcription termination site (TTS) confirmed that this relative increase in occupancy downstream of the TSS did not extend across gene bodies (**Sup. Fig. 5b**). The reduction in Pol II occupancy downstream of the TTS during HSV-1 infection reflects loss of Pol II pausing at the TTS associated with disruption of transcription termination previously reported in HSV-1 infection (8). Interestingly, genes with increased PI upon HSV-1 infection only showed a small change in Pol II occupancy in sense direction at the promoter in HSV-1 infection (**Fig. 2b**). In contrast, genes with strong PI reduction upon HSV-1 infection showed a strong reduction of the major peak height at the TSS, a pronounced broadening of the peak into the gene body and a second minor peak downstream the TSS (**Fig. 2c**). A similar but less pronounced effect was observed for genes with a weak reduction in PI, with a general broadening of the TSS peak but no minor peak (**Fig. 2d**). To compare this observation against "normal" reduction in pausing linked to increased elongation, we also performed metagene analyses of the PRO-seq data for CDK7 inhibition. This showed a reduction in the TSS peak height upon CDK7 inhibition, with little broadening of the peak into the gene body, but increased Pol II occupancy until the end of the promoter window (**Sup. Fig. 5c**). Thus, the widespread PI reduction in HSV-1 infection does not simply reflect reduction of Pol II pausing with increased elongation but more complex changes in Pol II occupancy.

To identify groups of genes with distinct changes of Pol II occupancy around the TSS between HSV-1 and mock infection, we performed hierarchical clustering of genes based on their PRO-seq profiles in sense direction for both mock and HSV-1 infection (**Fig. 3a**). Since we wanted to ensure that genes with distinct patterns of changes were placed in different clusters, a stringent cutoff was applied on the clustering dendrogram to obtain 50 gene clusters at the cost of obtaining multiple clusters with similar patterns. While most clusters exhibited only a narrow peak at the TSS in mock infection (e.g., **Fig. 3b**), some already exhibited a second minor peak shortly after the TSS already prior to infection (e.g., **Fig. 3c,d**). In addition, a few clusters representing a total of 2,018 genes showed peaks in mock infection that were shifted relative to the TSS we had identified (e.g., **Fig. 3d**, **Sup. Fig. 6a**). In all cases, the peak was shifted at most 750 bp from the identified TSS and was commonly within 100-200 bp of the identified TSS. Since position of peaks within clusters were highly similar due to the stringent clustering cutoff, analysis of individual clusters thus avoids confounding effects resulting from mis-identification of TSS positions. Typical Pol II occupancy changes upon HSV-1 infection included a reduction of peak height at the TSS with a broadening of the

peak into the gene body (e.g., **Fig. 3b**), changes in minor and major peak heights in case multiple peaks were already present prior to infection (e.g., **Fig. 3c,d**) as well as new peaks originating downstream of the TSS in HSV-1 infection (e.g., **Fig. 3e**). **Sup. Fig. 7** provides an overview on positions, number and relative heights of Pol II occupancy peaks in mock and HSV-1 infection for each cluster. In total, 30 clusters shared the same major TSS peak between mock and HSV-1 infection, which included also clusters with only a reduction in peak height but no broadening of the TSS peak (Cluster 20; **Sup. Fig 6b**) and clusters without loss of Pol II pausing (Clusters 21 & 27; **Sup. Fig 6c,d**). 28 clusters showed a second peak in HSV-1 infection downstream of the TSS peak with a median distance of 480 bp to the major peak (e.g., **Fig. 3c,d**). For 9 of these, the major TSS peak differed between mock and HSV-1 infection (e.g., **Fig. 3e**). Notably, almost all clusters exhibited a reduced Pol II peak height at the TSS, except Clusters 27 and 21 (262 genes; **Sup. Fig. 6c,d**). Read coverage plots for example genes from different clusters are shown in **Sup. Fig. 8**.

Since increasing or newly originating secondary PRO-seq peaks could also represent alternative transcription initiation, we next investigated the presence of alternative TSSs in either the PROcap-seq and PRO-seq data of flavopiridol-treated HFF or the genome annotation. For most clusters, <15% of genes showed evidence for an alternative TSS at the additional peak positions in either flavopiridol-treated cells (e.g., **Sup. Fig. 9a,b**) or the genome annotation (e.g., **Sup. Fig. 9d,e**). This is in clear contrast to clusters in which the TSS identified from flavopiridol-treated HFF did not represent the dominant TSS in HEp-2 cells. Here, almost 50% of genes had an additional peak in flavopiridol-treated cells or an annotated transcript start at the position of the dominant TSS in HEp-2 cells (e.g., **Sup. Fig. 9c,f**). We furthermore investigated induction of alternative *de novo* transcription initiation downstream of the TSS during HSV-1 infection using dRNA- and cRNA-seq data of transcript 5'ends for mock and HSV-1 infection from our recent re-annotation of the HSV-1 genome (n=2 replicates) (42). dRNA-seq is based on selective cloning and sequencing of the 5′-ends of cap-protected RNA molecules resistant to the 5′–3′-exonuclease XRN1. cRNA-seq is based on circularization of RNA fragments. Both methods strongly enrich reads from 5′ RNA ends. dRNA-seq was performed for mock and 8 h p.i. HSV-1 infection with and without XRN1 treatment. cRNA-seq was performed for mock, 1, 2, 4, 6 and 8 h p.i. HSV-1 infection. Metagene analyses of dRNA- and cRNA-seq data showed clear peaks coinciding with the major PRO-seq peaks in mock infection and smaller peaks at minor PRO-seq peaks already present in mock infection (**Sup. Fig. 10,11**). In contrast, no increase in transcript 5'ends was

observed at the positions of secondary PRO-seq peaks that increase or newly originate during HSV-1 infection (**Sup. Fig. 10,11**). In summary, changes in Pol II occupancy during HSV-1 infection are not due to alternative initiation at novel TSSs leading to capped transcripts, however we cannot completely exclude that they may reflect abortive *de novo* initiation at novel initiation sites downstream of the TSS. This analysis also excludes Pol II creeping as observed upon H2O2 treatment (43) as the latter would lead to signals from capped transcripts increasing downstream of the TSS in the pausing region.

Considering the bidirectionality of transcription initiation at human promoters, we also performed metagene analyses of PRO-seq profiles in antisense direction. Notably, antisense transcription initiation at bidirectional promoters commonly only results in short, unspliced, non-polyadenylated and unstable upstream antisense RNAs (uaRNAs) (44), that have highly heterogeneous 3′ ends (45). The metagene analysis on all genes of antisense PRO-seq profiles also showed a reduction in the antisense TSS peak height and a broadening of the peak in antisense direction (**Sup. Fig. 5a**). However, clustering of antisense PRO-seq profiles in mock and HSV-1 infection with the same approach as for sense profiles to obtain 50 clusters did not identify different patterns between clusters. Most of the 50 clusters exhibited only the same pattern as the metagene analysis of all genes (e.g., **Fig. 4a**). Only one cluster (430 genes) showed a small secondary antisense peak originating in HSV-1 infection (**Fig. 4b**), while another cluster (129 genes) showed a secondary peak that was already present in mock infection but increased relative to the TSS peak in WT infection (**Fig. 4c**).

HSV-1-induced increases of chromatin accessibility downstream of the TSS

We previously performed ATAC-seq to investigate chromatin accessibility for mock and WT-17 HSV-1 infection of HFFF at 2, 4, 6 and 8 h p.i. (n=2 replicates) (46). More recently, we also performed ATAC-seq for mock and WT-17 infection ± inhibition of viral DNA replication by phosphonoacetic acid (PAA), and infection with null mutants of the immediate early proteins ICP0, ICP4, ICP22 and ICP27 and the virion host shut-off protein (*vhs*) (n=2 replicates, 12 h p.i. for ΔICP0 and ΔICP22, 8 h p.i. for the rest) (47). In an independent experiment, ATAC-seq was also performed for mock and WT-F infection at 8 and 12 h p.i. (n=2) (47). We employed these data to investigate whether alterations in promoter-proximal Pol II pausing during HSV-1 infection were associated with changes in chromatin accessibility around promoters. Metagene analyses of all analyzed genes showed the expected

peak in chromatin accessibility directly upstream of the TSS and the sense PRO-seq peak in mock, WT-17 and WT-F infection, with a small minor peak downstream of the TSS (**Fig. 5a,b**, **Sup. Fig 12a**). Consistent with an overall reduction of Pol II occupancy at cellular promoters during HSV-1 infection (4, 5), the height of the major ATAC-seq peak decreased continuously with the duration of infection. Strikingly, however, accessibility at the small minor peak downstream of the TSS increased relative to the major peak, with changes clearly detectable from 4 h p.i. **(Fig. 5a,b**, **Sup. Fig 12a**). This effect was highly reproducible in the three independent experiments. Analysis of the 50 clusters defined on the sense PRO-seq profiles showed that changes in chromatin accessibility downstream of the TSS commonly matched the changes in Pol II occupancy observed in the PRO-seq data. Strong increases in chromatin accessibility at the minor peak were generally observed for clusters that showed a new or increased second PRO-seq peak or a strong broadening of the major PRO-seq peak downstream of the TSS (**Fig. 5d, 3c** for Cluster 4, **Suppl. Fig. 13 a,b,d,e** for Clusters 11 and 6). In some cases, we even observed a downstream shift of the major ATAC-seq peak (**Fig. 5e,3d** for Cluster 47). In contrast, clusters without pausing changes (**Fig. 5f**, **Sup. Fig. 6d** for Cluster 27) or only a reduction in the height of the PRO-seq peak (**Sup. Fig. 13c,f** for Cluster 20) showed no or very little increases in chromatin accessibility downstream of the TSS. Thus, downstream shifts of promoter-proximal pausing coincided with an extension of the nucleosome-free region at the TSS.

Comparison of ATAC-seq data of WT-17 and null mutant infections from the same experiment showed that chromatin accessibility at the minor TSS peak was increased in ΔICP4 infection similar to but less pronounced than in WT-17 (**Fig. 5c, Sup. Fig. 12b**). This trend was also observed when looking at individual clusters (**Fig. 5d-f**, **Sup. Fig. 13**). Here, differences between mock and ΔICP4 infection were statistically significant for the minor peak regions downstream of the TSS (Wilcoxon rank sum test per bin, multiple testing adjusted p-value $<10^{-3}$, **Sup. Fig. 14a,d,g,j)**). However, p-values were not as low as for the comparison between mock and WT-17 infection (**Sup. Fig. 14b,e,h,k)**) and WT-17 and ΔICP4 infection **(Sup. Fig. 14c,f,i,l**). PAA treatment during WT-17 infection showed a similar trend as ΔICP4 infection at the minor peak with less increase of chromatin accessibility than observed without PAA treatment (**Fig. 5c**). In contrast, the major peak height was comparable to WT-17 infection and more strongly reduced than in ΔICP4 infection. This suggests that the increase of chromatin accessibility downstream of the TSS is independent of the reduction directly at the TSS and depends on late gene expression as ICP4

is required for early and late expression and PAA inhibits late gene expression. Consistent with changes at the major and minor peak being independent, ΔICP22 infection resulted in a further reduction of the major peak height but was comparable to WT-17 infection at the minor downstream peak (**Sup. Fig. 12c**). For ΔICP0, ΔICP27 and Δ*vhs* infection, changes in chromatin accessibility were almost identical to WT-17 infection (**Sup. Fig. 12d-f**). Read coverage plots for example genes are shown in **Sup. Fig. 15**.

To quantify the changes in chromatin accessibility around the TSS for each cluster during HSV-1 and mutant virus infections, we determined the position of highest chromatin accessibility within ± 3kb of the TSS for each cluster in mock infection and then calculated the difference between metagene curves in HSV-1 infection compared to mock infection separately in the 720 bp up- and downstream of this position (**Sup. Fig. 16**). Differences upstream of the peak maximum were mostly negative (i.e., normalized chromatin accessibility was lower in HSV-1 infection), but always positive downstream of the peak maximum (i.e., normalized chromatin accessibility was higher in HSV-1 infection) and absolute differences were larger downstream of the peak maximum. Strong increases in chromatin accessibility downstream of the peak maximum were observed for 14 of the 50 cluster (Clusters 1, 4, 5, 11-14, 26, 31, 32, 43, 47, 49 and 50), all of which showed a strong reduction in the major PRO-seq peak height in HSV-1 infection. Almost all of these (13/14=93%) exhibited a novel or increasing secondary PRO-seq peak downstream of the TSS. The one exception, Cluster 13, showed a strong broadening of the PRO-seq peak into downstream regions. In contrast, only small differences were observed for clusters that showed no Pol II pausing changes in the PRO-seq profiles (Clusters 21 and 27), no broadening of the PRO-seq sense peak into downstream regions (Clusters 7 and 20) or a complete loss of the promoter peak in HSV-1 infection (Clusters 23 and 24). Across all clusters, WT-17 infection plus PAA treatment and ΔICP4 infection showed the smallest relative increases in chromatin accessibility downstream of the TSS, but increases were observed. In summary, our data show increased chromatin accessibility downstream of the TSS in HSV-1 infection for most genes that was alleviated but not abolished upon PAA treatment and in absence of ICP4 and was correlated to downstream shifts in promoter-proximal pausing during HSV-1 infection.

H2A.Z-containing +1 nucleosomes are shifted downstream in HSV-1 infection

Due to the distinct association of H2A.Z with -1 and +1 nucleosomes and its previously reported role in maintenance of Pol II pausing, we performed ChIPmentation of mock and 8 h

p.i. WT-17 infection of HFFF with an antibody recognizing the C-terminal part of H2A.Z (n=3 replicates). H2A.Z 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 (48). Recently, Mylonas *et al*. showed specifically for H2A.Z.1 that its depletion using an inducible protein degron system reduces PIs and that H2A.Z.1 slows the release of Pol II from pausing (32). 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 should recognize both isoforms. Metagene analyses for all analyzed genes showed the expected distribution of H2A.Z in mock infection with two equally high peaks on both sides of the TSS, corresponding to the -1 and +1 nucleosome positions (**Fig. 6a**). Strikingly, however, HSV-1 infection resulted in a significant redistribution of H2A.Z, with the +1 nucleosome peak increasing relative to the -1 nucleosome peak and shifting further into the gene body. Comparison to PRO-seq profiles showed that this mirrored largely the broadening of the PRO-seq TSS peak during HSV-1 infection (**Fig. 6b**). Differences in H2A.Z occupancy between mock and HSV-1 infection were highly significant across most of the promoter region in the metagene analysis of all analyzed genes (Wilcoxon rank sum test per bin, multiple testing adjusted p-value $<10^{-3}$, **Fig. 6a**). When analyzed separately for the 50 clusters, significance was largely limited to the +1 nucleosome peak region, the region from -3 to -2 kb of the TSS and sometimes the -1 nucleosome peak region (**Fig. 6c**, **Sup. Fig. 17a-f**). Although clusters without changes in PRO-seq profiles (Clusters 21 and 27) or only a reduction in PRO-seq peak height (Cluster 20) also showed small changes in the +1 nucleosome peak region, these were not statistically significant (**Sup. Fig. 17g-i**). Read coverage plots for example genes are shown in **Fig. 6d** and **Sup. Fig. 18**. To quantify the effect for all clusters, we performed a similar analysis as for the ATAC-seq data to determine the difference in H2A.Z occupancy profiles up- and downstream of the position of the minimum between the two H2A.Z peaks in mock infection. Here, almost all clusters showed a decrease upstream of this minimum (i.e, at -1 nucleosome positions) and an increase downstream of this minimum (i.e, at +1 nucleosome positions) during HSV-1 infection (**Sup. Fig. 19**).

Mylonas *et al*. also reported that H2A.Z.1 controls the dynamics of the pausing factor NELF (32). Interestingly, a recent study by Aoi *et al.* showed that rapid NELF depletion does not completely abolish pausing but Pol II is instead paused at a 2nd more downstream pause site

around the +1 nucleosome (49). Mylonas *et al.* proposed that H2A.Z.1 is necessary for this additional pause. Since Rivas *et al.* showed an ICP4-dependent decrease of NELF in the promoter-proximal region of some HSV-1-activated genes (9), we reanalyzed PRO-seq data from the Aoi *et al.* study for 0, 1, 2 and 4 h auxin-induced degradation of NELF to investigate potential similarities to HSV-1 infection. For a few of our clusters, we did indeed observe an increased second PRO-seq peak upon NELF degradation or a broadening of the first PRO-seq peak downstream of the TSS (e.g., **Sup. Fig. 20a-h**). For most clusters, however, we only observed a reduction in the major peak height and a minor broadening of the peak (e.g., **Sup. Fig. 20i-n**). In either case, the alterations in the PRO-seq profiles in HSV-1 infection were much more pronounced than after NELF depletion, with more extensive broadening of peaks and new secondary peaks being further downstream of the major peak. In summary, delayed pause sites in HSV-1 infection are further downstream than "normal" 2nd pause sites at H2A.Z-containing +1 nucleosome positions used upon NELF depletion but appear to be associated with the shifted +1 nucleosome positions in HSV-1 infected cells.

Discussion

Promoter-proximal Pol II pausing is a key regulatory step between transcription initiation and productive elongation. HSV-1 infection dramatically impacts Pol II positioning on host genes, including promoter-proximal regions (5), and promoter-proximal Pol II pausing on HSV-1 genes plays a key role in HSV-1 transcription (50). Rivas *et al.* recently reported that HSV-1 infection leads to a reduction in pausing indices for activated host genes and an increase in pausing indices at repressed host genes (9). Our re-analysis of PRO-seq data of mock and WT-F infection revealed a widespread reduction in pausing indices for most expressed host genes. Since pausing indices are an imperfect though easy-to-use measure of promoter-proximal Pol II pausing and are changed by various types of alterations in Pol II pausing, we performed metagene analyses to further characterize alterations in Pol II occupancy around promoters during HSV-1 infection. This revealed that HSV-1 infection does not lead to a straightforward reduction of promoter-proximal Pol II pausing with a relative increase of elongating Pol II on the whole gene body. Instead, we observed that Pol II pausing is retained for the vast majority of genes but is shifted to downstream pause sites. This is reflected in broadened Pol II promoter-proximal peaks that extend further into the gene body and newly originating or increasing downstream peaks. A fine-grained clustering analysis identified a wide range of different patterns for different genes in HSV-1 infection, which contrasts with the sharp promoter-proximal peaks commonly observed in uninfected cells. This indicates

that positioning of shifted pause sites in HSV-1 infection is less well-defined than of “normal” pause sites in uninfected cells. Analysis of transcript start site profiling in HSV-1 infection excluded that this was due to *de novo* initiation at downstream sites.

Interestingly, analysis of promoter-proximal pausing in antisense direction also showed a broadening of antisense TSS peaks, however secondary antisense peaks were only observed for a small fraction of genes (7.3%). It has been proposed that Pol II is particularly prone to pausing and termination during early elongation, specifically on AT-rich sequences often found upstream of promoters (51). Previously, we reported both widespread disruption of transcription termination in HSV-1 infection (8) as well as widespread activation of antisense transcription at promoters and within gene bodies (52). It is thus tempting to speculate that activation of antisense transcription in HSV-1 infection could be linked to the alterations in antisense Pol II pausing observed in this study, potentially in combination with disrupted transcription termination.

Nucleosomes represent a natural barrier to transcription and are disassembled before and reassembled after transcribing Pol II (53). Nucleosomes directly downstream of the TSS are generally well-positioned at specific locations, in particular the +1 nucleosome, but less so further up- or downstream (54-56). Pol II pausing occurs between the promoter and the +1 nucleosome, and the +1 nucleosome has been shown to play a role in pausing, with strong positioning of the +1 nucleosome increasing pausing (57). Moreover, while NELF has previously been considered to be required for establishing Pol II pausing, rapid depletion of NELF using auxin-inducible degron does not abolish pausing (49). Instead, pausing appears to be a two-step process with Pol II transitioning from the 1st to a 2nd pause site upon NELF depletion. This 2nd pause site is associated with +1 nucleosomes and has been proposed to require the noncanonical histone variant H2A.Z enriched at +1 nucleosomes (32). Strikingly, investigation of nucleosome-free regions between the -1 and +1 nucleosomes prior and during HSV-1 infection using ATAC-seq showed an extension of the nucleosome-free region into downstream regions in particular for genes with prominent downstream shifts in pausing. Consistent with this, positioning of H2A.Z was altered during HSV-1 infection with the +1 nucleosome peak shifted further into the gene body. Our results indicate that altered positioning of the H2A.Z-containing +1 nucleosomes is linked to the changes in Pol II pausing upon HSV-1 infection.

Rivas *et al.* found that ICP4 was required both for the reduction of pausing indices at HSV-1-activated genes as well as for decreased levels of NELF at promoter regions of four activated genes tested by ChIP. In contrast, we found nucleosome-free regions were also extended downstream of the TSS upon ΔICP4 infection, albeit not as pronounced as in WT infection. This suggests that ICP4 is not generally required for downstream shifts in +1 nucleosomes in HSV-1 infection, but it may still contribute to the delay on promoter-proximal pausing, e.g., via its impact on NELF binding to promoters. Since depletion of NELF leads to pausing at the 2nd more downstream pause sites, depletion of NELF from host promoters via ICP4-dependent and/or ICP4-independent pathways in HSV-1 infection could partly explain the delay in Pol II pausing. The role of ICP22 in shaping Pol II pausing during WT HSV-1 infection also remains an open and intriguing question, considering that ectopic expression of an ICP22 segment interacting with CDK9 (20) does not delay but rather increases Pol II pausing similar to CDK9 inhibition (21). In any case, our data shows that ICP22 is not required for the extension of the nucleosome-free region at the promoter. It should be noted that NELF depletion alone did not lead to such an extensive delay of Pol II pausing with increasing or newly originating secondary Pol II peaks as observed upon HSV-1 infection. However, 2nd pause sites identified upon NELF depletion were associated with +1 nucleosomes and our analysis of H2A.Z positioning indicated a downstream shift of +1 nucleosomes. Thus, our data suggest a model in which loss of Pol II pausing at 1st pause sites combined with altered +1 nucleosome positioning in HSV-1 infection leads to delayed pausing at shifted 2nd pause sites.

Interestingly, polymerase loss has previously been shown to have an impact on nucleosome positioning in yeast, resulting in downstream shifts of coding-region nucleosomes, in particular +1 nucleosomes (58). Thus, downstream shifts of +1 nucleosome in HSV-1 infection could be a consequence of the depletion of Pol II from cellular genes and loss of host transcriptional activity. This is consistent with our observation that the downstream extension of nucleosome-free regions at the TSS was reduced in ΔICP4 infection and upon PAA treatment as both PAA treatment and ICP4 knockdown reduce the extensive loss of host transcriptional activity during HSV-1 infection (9, 47). Notably, -1 nucleosomes did not show a directional shift in the yeast study, which would be consistent with our observation that the downstream shift upon HSV-1 infection was observed only for the +1 nucleosome-associated H2A.Z peak and not the -1 nucleosome-associated peak. Although PRO-seq in WT-F infection was performed relatively early at 3 h p.i. when loss of host transcriptional activity is

less pronounced than at 8 h p.i., we recently showed that ~50% of transcriptional activity is already lost by 3 h p.i. (7). The fact that secondary PRO-seq peaks were frequently observed in sense but rarely in antisense direction further argues for two distinct mechanisms impacting Pol II pausing in HSV-1 infection. The first one would then reduce pausing at 1st pause sites and affect both sense and antisense transcription. In contrast, the second one would impact positioning of +1 nucleosomes and corresponding 2nd pause sites and thus only sense transcription.

Materials and Methods

Previously published sequencing data analyzed in this study

PROcap-seq and PRO-seq data of flavopiridol-treated uninfected HFF cells were taken from the study by Parida *et al.* (35) (GEO accession: GSE113394, samples GSM3104917 and GSM3104913). PRO-seq data for mock and WT-F infection at 3 h p.i. of HEp-2 cells were taken from the study by Birkenheuer *et al.* (5) (n=3 replicates, GEO accession: GSE106126, samples GSM2830123 - GSM2830127). PRO-seq data for 0, 1, 2 and 4 h auxin-induced degradation of NELF were taken from the study by Aoi *et al.* (49) (n=1 apart from 0 h with n=2, GEO accession: GSE144786, samples GSM4296314 - GSM4296316, GSM4296318, GSM4296319). PRO-seq data of 15 min, 30 min, 1 h and 2 h of THZ1 treatment and 1 h DMSO treatment of Kasumi-1 cells were taken from the study by Sampathi *et al.* (39) (n=2, GEO accession: GSE102243, samples GSM2731702 - GSM2731711).

4sU-seq data for mock and hourly intervals for the first 8 h of WT-17 infection of HFFF was taken from our previous study (8) (n=2, GEO accession: GSE59717). dRNA-seq for mock and 8 h p.i. HSV-1 infection with and without XRN1 treatment and and cRNA-seq data for mock, 1, 2, 4, 6 and 8 h p.i. HSV-1 infection were taken from our previous study (42) (n=2, GSE128324, samples GSM3671394 - GSM3671411).

ATAC-seq data for mock and WT-17 HSV-1 infection of HFFF at 2, 4, 6 and 8 h p.i. (n=2 replicates) was taken from our previous publication (46) (GEO accession: GSE100611). ATAC-seq for mock and WT-17 infection (8 h p.i.), WT-17 infection with 8 h PAA treatment, infection with ICP0-null mutant (ΔICP0, strain 17, (59), 12 h p.i.), ICP22-null mutant (ΔICP22, R325, strain F, (60), 12 h p.i.), ICP27-null mutant (ΔICP27, strain KOS, (61), 8 h p.i.) and *vhs*-null mutant (Δvhs, strain 17, (62)) were taken from (47) (n=2 apart from ΔICP22 infection with n=4, GEO accession: GSE185234). This experiment also

included infection with an ICP4-null mutant (ΔICP4, n12, strain KOS, (63), 8 h p.i.), which had not previously been published (GEO accession: GSE212820). ATAC-seq data for an independent experiment for mock and WT-F infection at 8 and 12 h p.i. (n=2) was also taken from (47) (GEO Accession: GSE185235).

ChIPmentation, library preparation and sequencing

Two days prior to infection, two million HFFF 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 N2.

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 H2O 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 uL 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 H2O.

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) (see Supplementary Table 3); 3.75 μL H20 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. ChIPmentation data are available under GEO accession GSE212867.

Read alignment

The read alignment pipeline was implemented and run in the workflow management system Watchdog (64, 65). Published sequencing data were first downloaded from SRA using the sratoolkit version 2.10.8. Sequencing reads were aligned against the human genome (GRCh37/hg19) and human rRNA sequences using ContextMap2 version 2.7.9 (66) (using

BWA as short read aligner (67) and allowing a maximum indel size of 3 and at most 5 mismatches). For sequencing data of HSV-1 infection, alignment also included the HSV-1 genome (Human herpesvirus 1 strain 17, GenBank accession code: JN555585). 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 (68). Read coverage in bedGraph format was calculated from BAM files using BEDTools (69).

Data plotting and statistical analysis

All figures were created in R and all statistical analyses were performed in R (70). Read coverage plots were created using the R Bioconductor package Gviz (71). Differential gene expression analysis was performed using the R Bioconductor package DESeq2 (72) on gene read counts determined with using featureCounts (73) and gene annotations from Ensembl (version 87 for GRCh37) (74) on 4sU-seq data in a strand-specific manner. All fragments (read pairs for paired-end sequencing or reads for single-end sequencing) overlapping exonic regions on the corresponding strand by ≥25bp were counted for the corresponding gene. P-values for differential gene expression analysis were adjusted for multiple testing using the method by Benjamini and Hochberg (75)

Transcription start site identification

We used the iTiSS program to identify candidate TSS in PROcap-seq and PRO-seq of flavopiridol-treated HFF (42, 76). For this purpose, iTiSS was run separately for each sample in the SPARSE_PEAK mode with standard parameters. Afterwards, the iTiSS TSRMerger program was used to select only peaks that were identified in both samples within +/- 5 bp. Consistent peaks were only further considered if they were within 500 bp to the nearest annotated gene and for each gene the TSS with the highest read count (weighted by the number of possible alignments for the read) was selected for further analyses.

Calculation of pausing indices

PRO-seq read counts in promoter windows (TSS to TSS + 250 bp) and gene bodies (TSS + 250 bp to TSS + 2,250 bp or gene 3’ end if closer) were determined using featureCounts (73) and gene annotations from Ensembl (version 87 for GRCh37) (74) in a strand-specific manner and normalized by the total number of reads and window lengths to obtain RPKM values. RPKM values were averaged between replicates and genes with zero reads in either promoter

or gene body window were excluded from analysis. PI for a gene was then calculated as the ratio of promoter RPKM to gene body RPKM.

Metagene and clustering analysis

Metagene analyses were performed as previously described (77) 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 bp bins for each gene. For each bin, the average coverage per genome position was calculated (in a strand-specific manner for PRO-seq, non-strand-specific for ATAC-seq and ChIPmentation) 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 and metagene plots then show the average metagene curves across replicates. Genes without any reads in any of the analyzed samples were excluded from the analysis. For metagene analyses on the whole gene, the regions from -3 kb to +1.5 kb of the TSS and from -1.5 kb to +3 kb of the TTS were divided into 90 bp bins and the remainder of the gene body (+1.5 kb of TSS to −1.5 kb of TTS) into 100 bins of variable length in order to compare genes with different lengths. Genes with a gene length < 3 kb were excluded as regions around the TSS and TTS would overlap otherwise. To determine statistical significance of differences between average metagene curves for two conditions, paired Wilcoxon signed rank tests were performed for each bin comparing normalized coverage values for each gene for this bin between the two conditions. P-values were adjusted for multiple testing with the Bonferroni method across all bins within each subfigure and are color-coded in the bottom track of subfigures: red = adj. p-value $\leq 10^{-15}$, orange = adj. p-value $\leq 10^{-10}$, yellow = adj. p-value $\leq 10^{-3}$.

For hierarchical clustering analysis, PRO-seq profiles for each gene and condition were calculated for sense or antisense strand as for metagene analyses (without averaging across genes). PRO-seq profiles in promoter windows for mock and WT-F infection at 3 h p.i. were then concatenated and divided by the maximum value in the concatenated vector. Hierarchical clustering was performed in using the hclust function in R according to Euclidean distances and Ward's clustering criterion. Peaks in metagene plots for each cluster were then determined in the following way: First, all local and global maxima and minima of metagene curves for each condition were identified for each cluster using the find_peaks function in the

R ggpmisc package. The major peak was the global maximum. Subsequently, the next highest local maxima up- or downstream of the major peak were determined and retained as secondary peaks if (i) they were sufficiently removed from the borders of the 6 kb promoter window (i.e., within bins 30 to 80 of the 101 bins), (ii) the difference between the height of the secondary peak and the minimum value between the major and secondary peak was at least 10% of the major peak height and (iii) the height of the secondary peak was at least 20% of the major peak height.

Code availability

Workflows for PI calculation, metagene analyses, clustering and figure creation were implemented and run in Watchdog (64, 65) and are available at https://doi.org/10.5281/zenodo.7041444. Corresponding Watchdog modules are available in the Watchdog module repository (https://github.com/watchdog-wms/watchdog-wms-modules/).

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the framework of the Research Unit FOR5200 DEEP-DV (443644894) project FR 2938/11-1 and by grants FR2938/9-1 to C.C.F. and LD1275/6-1 to L.D.

Competing interests

The authors declare no competing interests.

References

1. 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.
2. Oroskar AA RG. Control of mRNA stability by the virion host shutoff function of herpes simplex virus. J Virol. 1989;63:1897–906.
3. 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.
4. 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.

5. 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).

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

7. 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).

8. 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.

9. Rivas T, Goodrich JA, Kugel JF. The herpes simplex virus 1 protein ICP4 acts as both an activator and repressor of host genome transcription during infection. Mol Cell Biol. 2021:MCB0017121.

10. Gilmour DS, Lis JT. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. Mol Cell Biol. 1986;6(11):3984-9.

11. Rougvie AE, Lis JT. The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell. 1988;54(6):795-804.

12. Landick R. The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans. 2006;34(Pt 6):1062-6.

13. Sheridan RM, Fong N, D’Alessandro A, Bentley DL. Widespread Backtracking by RNA Pol II Is a Major Effector of Gene Activation, 5′ Pause Release, Termination, and Transcription Elongation Rate. Molecular Cell. 2019;73(1):107-18.e4.

14. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet. 2012;13(10):720-31.

15. Yamaguchi Y, Shibata H, Handa H. Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. Biochim Biophys Acta. 2013;1829(1):98-104.

16. Marshall NF, Price DH. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J Biol Chem. 1995;270(21):12335-8.

17. Vos SM, Farnung L, Boehning M, Wigge C, Linden A, Urlaub H, et al. Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature. 2018;560(7720):607-12.

18. Zhou Q, Li T, Price DH. RNA polymerase II elongation control. Annu Rev Biochem. 2012;81:119-43.

19. Laitem C, Zaborowska J, Isa NF, Kufs J, Dienstbier M, Murphy S. CDK9 inhibitors define elongation checkpoints at both ends of RNA polymerase II–transcribed genes. Nature Structural & Molecular Biology. 2015;22(5):396-403.

20. Zaborowska J, Baumli S, Laitem C, O'Reilly D, Thomas PH, O'Hare P, et al. Herpes Simplex Virus 1 (HSV-1) ICP22 Protein Directly Interacts with Cyclin-Dependent Kinase (CDK)9 to Inhibit RNA Polymerase II Transcription Elongation. PLOS ONE. 2014;9(9):e107654.

21. 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.

22. Watson RJ, Clements JB. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. Nature. 1980;285(5763):329-30.

23. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The antisense transcriptomes of human cells. Science. 2008;322(5909):1855-7.

24. Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, et al. RNA exosome depletion reveals transcription upstream of active human promoters. Science. 2008;322(5909):1851-4.

25. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, et al. Divergent transcription from active promoters. Science. 2008;322(5909):1849-51.

26. Seila AC, Core LJ, Lis JT, Sharp PA. Divergent transcription: a new feature of active promoters. Cell Cycle. 2009;8(16):2557-64.

27. Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, et al. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell. 2010;143(4):540-51.

28. Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, et al. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev. 2008;22(14):1921-33.

29. 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.

30. 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.

31. Day DS, Zhang B, Stevens SM, Ferrari F, Larschan EN, Park PJ, et al. Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. Genome Biology. 2016;17(1):120.

32. Mylonas C, Lee C, Auld AL, Cisse II, Boyer LA. A dual role for H2A.Z.1 in modulating the dynamics of RNA polymerase II initiation and elongation. Nature Structural & Molecular Biology. 2021;28(5):435-42.

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

34. Conn KL, Hendzel MJ, Schang LM. Histone H2A and H2A.Z are mobilized in HSV-1 infected cells. 2013.

35. Parida M, Nilson KA, Li M, Ball CB, Fuchs HA, Lawson CK, et al. Nucleotide Resolution Comparison of Transcription of Human Cytomegalovirus and Host Genomes Reveals Universal Use of RNA Polymerase II Elongation Control Driven by Dissimilar Core Promoter Elements. mBio. 2019;10(1):e02047-18.

36. Williams Lucy H, Fromm G, Gokey Nolan G, Henriques T, Muse Ginger W, Burkholder A, et al. Pausing of RNA Polymerase II Regulates Mammalian Developmental Potential through Control of Signaling Networks. Molecular Cell. 2015;58(2):311-22.

37. Jonkers I, Kwak H, Lis JT. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. eLife. 2014;3:e02407.

38. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science. 2008;322(5909):1845-8.

39. Sampathi S, Acharya P, Zhao Y, Wang J, Stengel KR, Liu Q, et al. The CDK7 inhibitor THZ1 alters RNA polymerase dynamics at the 5' and 3' ends of genes. Nucleic Acids Res. 2019;47(8):3921-36.

40. Hutterer C, Eickhoff J, Milbradt J, Korn K, Zeittrager I, Bahsi H, et al. A novel CDK7 inhibitor of the Pyrazolotriazine class exerts broad-spectrum antiviral activity at nanomolar concentrations. Antimicrob Agents Chemother. 2015;59(4):2062-71.

41. Schang LM, Bantly A, Knockaert M, Shaheen F, Meijer L, Malim MH, et al. Pharmacological cyclin-dependent kinase inhibitors inhibit replication of wild-type and drug-resistant strains of herpes simplex virus and human immunodeficiency virus type 1 by targeting cellular, not viral, proteins. J Virol. 2002;76(15):7874-82.

42. Whisnant AW, Jurges 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.

43. Nilson KA, Lawson CK, Mullen NJ, Ball CB, Spector BM, Meier JL, et al. Oxidative stress rapidly stabilizes promoter-proximal paused Pol II across the human genome. Nucleic Acids Research. 2017;45(19):11088-105.

44. Lykke-Andersen S, Žumer K, Molska EŠ, Rouvière JO, Wu G, Demel C, et al. Integrator is a genome-wide attenuator of non-productive transcription. Molecular Cell. 2021;81(3):514-29.e6.

45. Scruggs Benjamin S, Gilchrist Daniel A, Nechaev S, Muse Ginger W, Burkholder A, Fargo David C, et al. Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. Molecular Cell. 2015;58(6):1101-12.

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

47. Djakovic L, Hennig T, Reinisch K, Milic A, Whisnant A, Wolf K, et al. The HSV-1 ICP22 protein selectively impairs histone repositioning upon Pol II transcription downstream of genes. Research Square; 2021.

48. 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.

49. Aoi Y, Smith ER, Shah AP, Rendleman EJ, Marshall SA, Woodfin AR, et al. NELF Regulates a Promoter-Proximal Step Distinct from RNA Pol II Pause-Release. Molecular Cell. 2020;78(2):261-74.e5.

50. Birkenheuer CH, Baines JD. RNA Polymerase II Promoter-Proximal Pausing and Release to Elongation Are Key Steps Regulating Herpes Simplex Virus 1 Transcription. Journal of Virology. 2020;94(5):e02035-19.

51. Vlaming H, Mimoso CA, Field AR, Martin BJE, Adelman K. Screening thousands of transcribed coding and non-coding regions reveals sequence determinants of RNA polymerase II elongation potential. Nature Structural & Molecular Biology. 2022;29(6):613-20.

52. 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.

53. Petesch SJ, Lis JT. Overcoming the nucleosome barrier during transcript elongation. Trends in Genetics. 2012;28(6):285-94.

54. Bai L, Morozov AV. Gene regulation by nucleosome positioning. Trends Genet. 2010;26(11):476-83.

55. Schones DE, Cui K, Cuddapah S, Roh T-Y, Barski A, Wang Z, et al. Dynamic Regulation of Nucleosome Positioning in the Human Genome. Cell. 2008;132(5):887-98.

56. Struhl K, Segal E. Determinants of nucleosome positioning. Nature Structural & Molecular Biology. 2013;20(3):267-73.

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

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

59. 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.

60. 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.

61. 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.

62. 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.

63. DeLuca NA, Schaffer PA. Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. Journal of Virology. 1988;62(3):732-43.

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

65. Kluge M, Friedl M-S, Menzel AL, Friedel CC. Watchdog 2.0: New developments for reusability, reproducibility, and workflow execution. GigaScience. 2020;9(6).

66. 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.

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

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

69. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.
70. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria2022.
71. Hahne F, Ivanek R. Visualizing Genomic Data Using Gviz and Bioconductor. In: Mathé E, Davis S, editors. Statistical Genomics: Methods and Protocols. New York, NY: Springer New York; 2016. p. 335-51.
72. 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.
73. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2013;30(7):923-30.
74. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode M R, Armean Irina M, et al. Ensembl 2022. Nucleic Acids Research. 2021;50(D1):D988-D95.
75. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B. 1995;57(1):289–300.
76. Jürges CS, Dölken L, Erhard F. Integrative transcription start site identification with iTiSS. Bioinformatics. 2021;37(18):3056-7.
77. 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 reports. 2019;20(9):e47592.

Figure legends

Fig. 1: HSV-1 infection impacts promoter-proximal pausing of most host genes.
(a) Scatter plots comparing pausing indices (PI) between mock and WT-F infection at 3h p.i. The dashed line indicates equal PI values and solid lines a 2-fold change in PIs. Genes were divided into three groups according to changes in their PI: (1) increased PI in HSV-1 infection (fold-change >1, 763 genes, red), (2) slightly reduced PI in HSV-1 infection (fold-change < 1 but ≥ 0.5, 2,082 genes, blue), (3) strongly reduced PI in HSV-1 infection (fold-change < 0.5, 4,211 genes, green). **(b,c)** Boxplots showing the distribution of **(b)** PIs and **(c)** gene body RPKM in mock and WT-F infection at 3h p.i. for the three gene groups defined in **(a)**. Boxes represent the range between the first and third quartiles for each condition. Black horizontal lines in boxes show the median. The ends of the whiskers (vertical lines) extend the box by 1.5 times the inter-quartile range. Data points outside this range (outliers) are shown as small

circles. Gene body RPKM for each gene was calculated on the region from +250 to +2,250 bp (or up to the gene 3'end if closer) from the TSS to exclude the promoter peak. **(d)** Boxplots showing the distribution of log2 fold-changes in 4sU-RNA for the first 8 h of WT-17 infection compared to mock for the three gene groups defined in **(a)**.

Fig 2: HSV-1 infection leads to a downstream shift in promoter-proximal Pol II pausing. **(a)** Read coverage around the TSS in PRO-Seq data (sense strand only) for mock (green) and WT-F infection at 3 h p.i. (blue) for example genes with a reduction in PI upon HSV-1 infection. Read coverage was normalized to total number of mapped reads and averaged between replicates. The identified TSS used in the analysis is indicated by a short vertical line below each read coverage track. Gene annotation is indicated at the top. Boxes represent exons and lines introns and direction is indicated by arrowheads. Genomic coordinates are shown on the bottom. Please note that figures are not centered around the TSS, but a larger region downstream of the TSS was included than upstream of the TSS. **(b-d)** Metagene plots showing the distribution of PRO-seq profiles in sense (dark green and blue) and antisense (light green and blue) direction from -3 kb to + 3 kb around the TSS for the three gene groups defined in **Fig. 1a** with increased PI **(b)**, strongly reduced PI **(c)**, and slightly reduced PI **(d)**. Mock infection is shown in dark (sense) and light green (antisense) and WT-F infection at 3 h p.i. in dark (sense) and light blue (antisense). For metagene plots, the TSS ± 3 kb promoter window for each gene was divided into 101 bins, PRO-seq read counts for each sample were determined for each bin, normalized to sequencing depth and averaged across replicates. Subsequently, bin values for each gene were normalized to sum up to 1 to obtain the Pol II occupancy profile in the promoter window. PRO-seq profiles were determined separately for sense and antisense strand.

Fig. 3: Distinct patterns of changes in promoter-proximal pausing upon HSV-1 infection. (a) Heatmap showing the result of the hierarchical clustering analysis of PRO-seq profiles in mock and WT-F infection. For clustering, PRO-seq profiles for mock and WT-F infection were first concatenated and then normalized by dividing by the maximum value in the concatenated profiles. 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 50 clusters (marked by colored rectangles between the dendrogram and heatmap). Clusters are numbered from top to bottom. **(b-e)** Metagene plots of

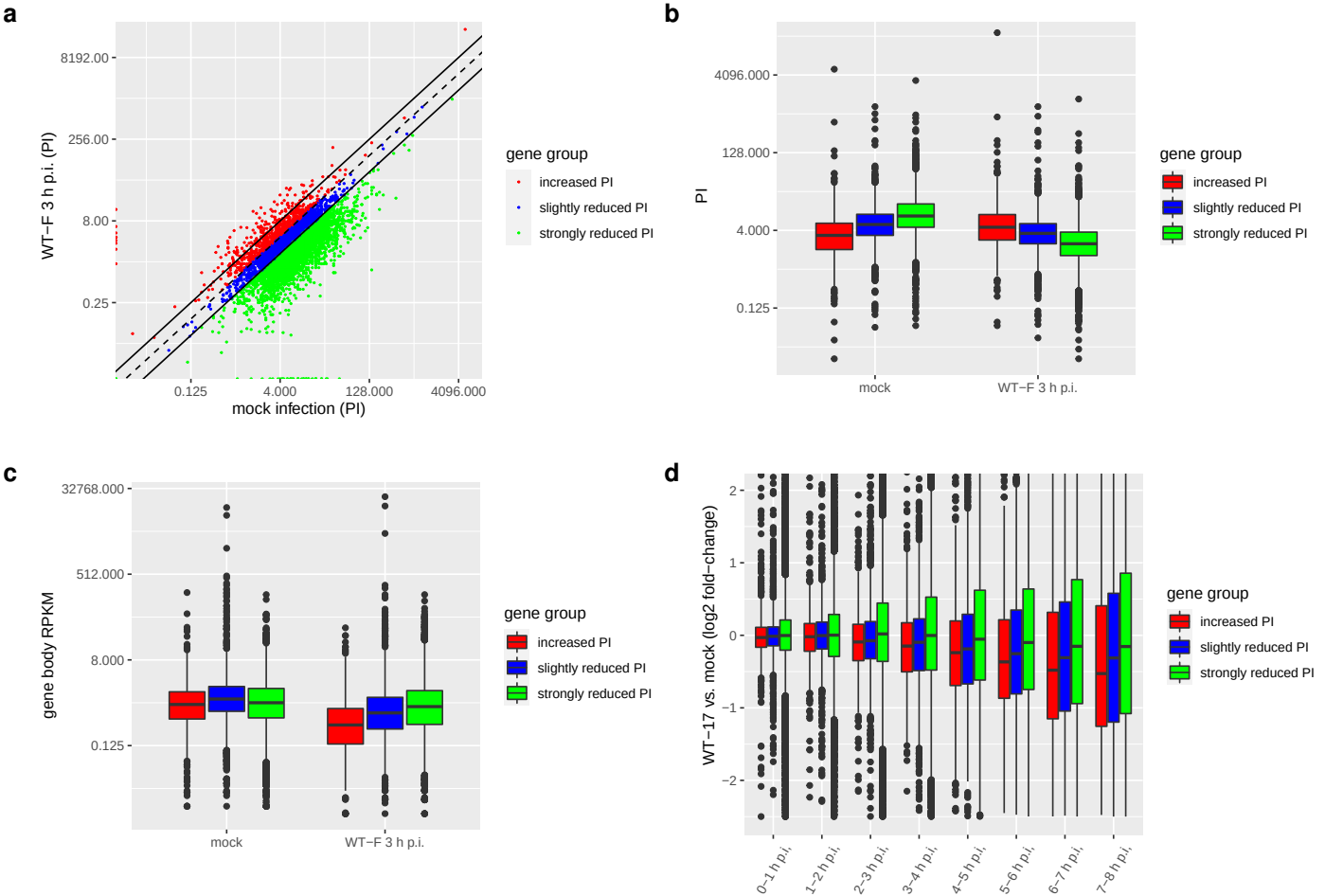
PRO-seq profiles on the sense strand in mock (dark green) and WT-F infection at 3 h p.i. (dark blue) for example Clusters 6, 4, 47 and 9 from **(a)**. See methods and legend to **Fig. 2** for an explanation of metagene plots.

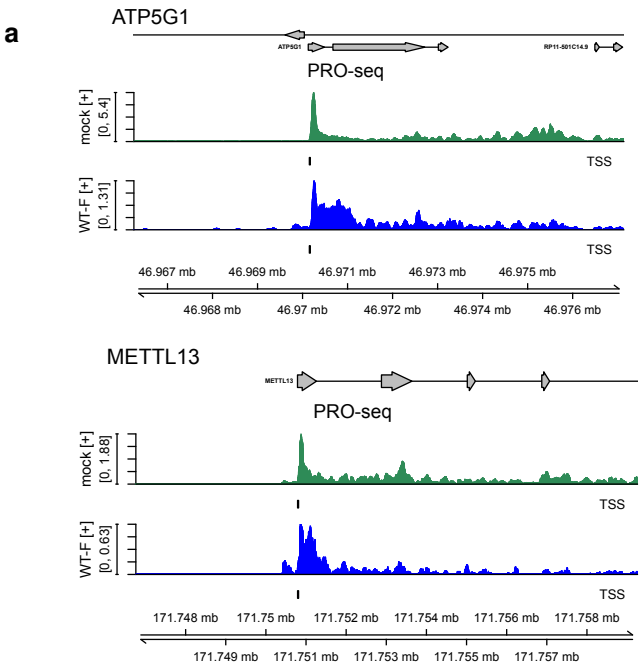
Fig. 4: Changes in antisense promoter-proximal pausing in HSV-1 infection. (a-c) Metagene plots of PRO-seq profiles on both sense (dark green and blue) and antisense (light green and blue) direction in mock (dark and light greeen) and WT-F infection at 3 h p.i. (dark and light blue)for example clusters resulting from the hierarchical clustering of genes according to antisense PRO-seq profiles in mock and WT-F infection. Here, clustering was performed as described in the legend for **Fig. 3a** but applied to concatenated antisense PRO-seq profiles in mock and WT-F infection. Thus, clusters shown here differ from the clusters shown in all other figures. **(a)** shows the most common pattern observed for almost all clusters with a broadening of the antisense PRO-seq peak at the TSS, while **(b)** and **(c)** show the only two clusters that exhibit different patterns with additional peaks originating or increasing in antisense direction during infection, respectively.

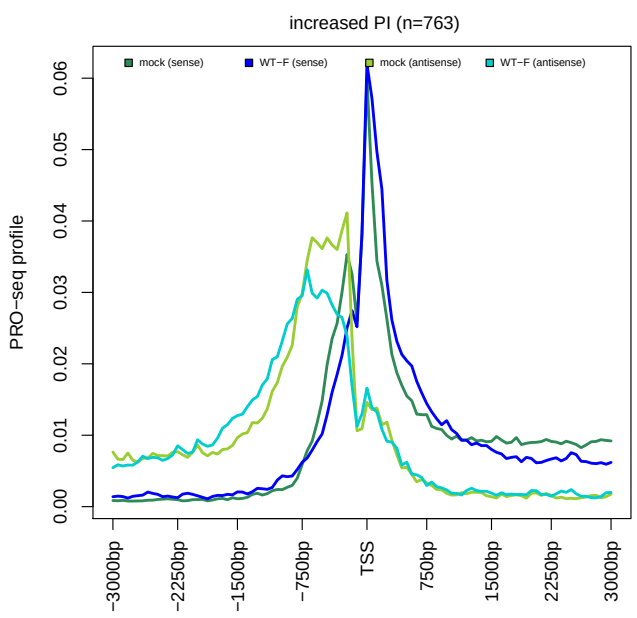
Fig. 5: HSV-1 infection leads to an extension of the nucleosome-free region around promoters into the gene body.
Metagene plots showing **(a)** ATAC-seq profiles for mock, 1, 2, 4, 6 and 8 h p.i. WT-17 infection and PRO-seq profiles on the sense strand in mock and WT-F 3 h p.i. infection, **(b)** ATAC-seq profiles for mock, 2, 4 and 8 p.i. WT-17 infection (a subset of time-points from the time-course shown in **(a)** shown for clarity), **(c)** ATAC-seq profiles for mock, WT-17 ± PAA and ΔICP4 infection at 8 h p.i. for all analyzed genes. ATAC-seq data were taken from our ATAC-seq experiment including WT-17 ± PAA treatment and null mutant infections (47). **(d-f)** ATAC-seq profiles for mock, WT-17 and ΔICP4 infection at 8 h p.i. for selected clusters from the clustering shown in **Fig. 3a**. ATAC-seq profiles were calculated from ATAC-seq data in the same way as PRO-seq profiles but without strand specificity.

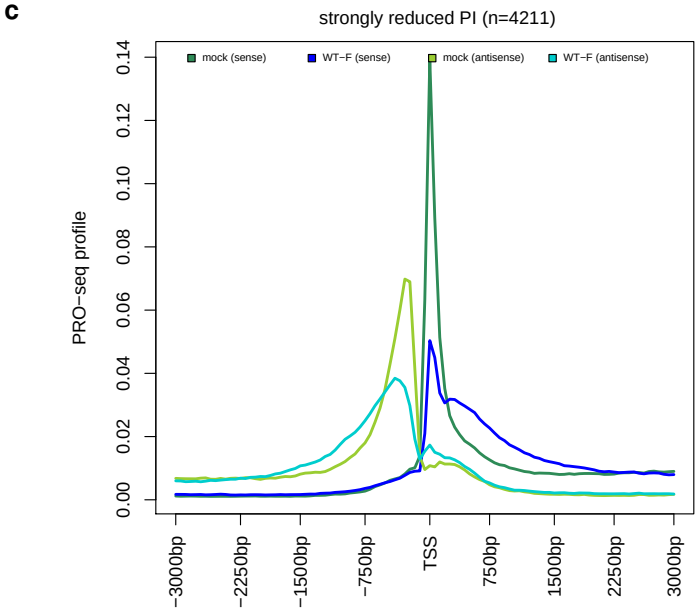
Fig. 6: HSV-1 infection leads to a downstream shift of H2A.Z enriched at +1 nucleosome positions. (a-c) Metagene plots showing H2A.Z occupancy profiles in mock and WT-17

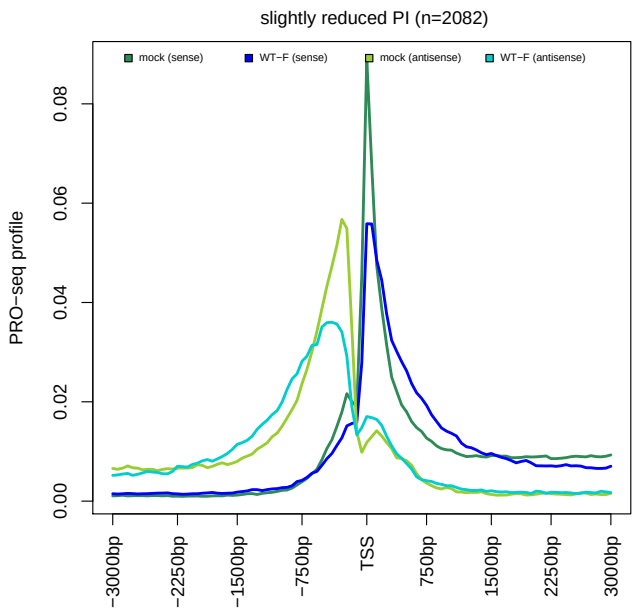
infection at 8 p.i. for **(a,b)** all analyzed genes and **(c)** an example cluster from the clustering shown in **Fig. 3a**. **(b)** shows the H2A.Z occupancy profiles from **(a)** together with the PRO-seq profiles on the sense strand for mock and WT-F infection at 3 h p.i. The color track at the bottom of subfigures **(a,c)** indicates the significance of paired Wilcoxon tests comparing the normalized H2A.Z coverage of genes for each bin between mock and WT-17 infection. P-values are adjusted for multiple testing with the Bonferroni method within each subfigure; color code: red = adj. p-value $\leq 10^{-15}$, orange = adj. p-value $\leq 10^{-10}$, yellow = adj. p-value $\leq 10^{-3}$. Metagene plots for further clusters are shown in **Sup. Fig. 17**. **(d)** Read coverage around the TSS in PRO-seq (sense strand only) and H2A.Z ChIPmentation (unstranded) data for mock (dark and light green) and HSV-1 infection (dark blue: WT-F 3 h p.i. and violet: WT-17 8 h p.i.) for the example gene ATP5G1. Figures for other example genes are shown in **Sup. Fig. 18**. Read coverage was normalized to total number of mapped reads and averaged between replicates. The identified TSS used in the analysis is indicated by a short vertical line below each PRO-seq read coverage track. Gene annotation is indicated at the top. Boxes represent exons and lines introns and direction is indicated by arrowheads. Genomic coordinates are shown on the bottom.

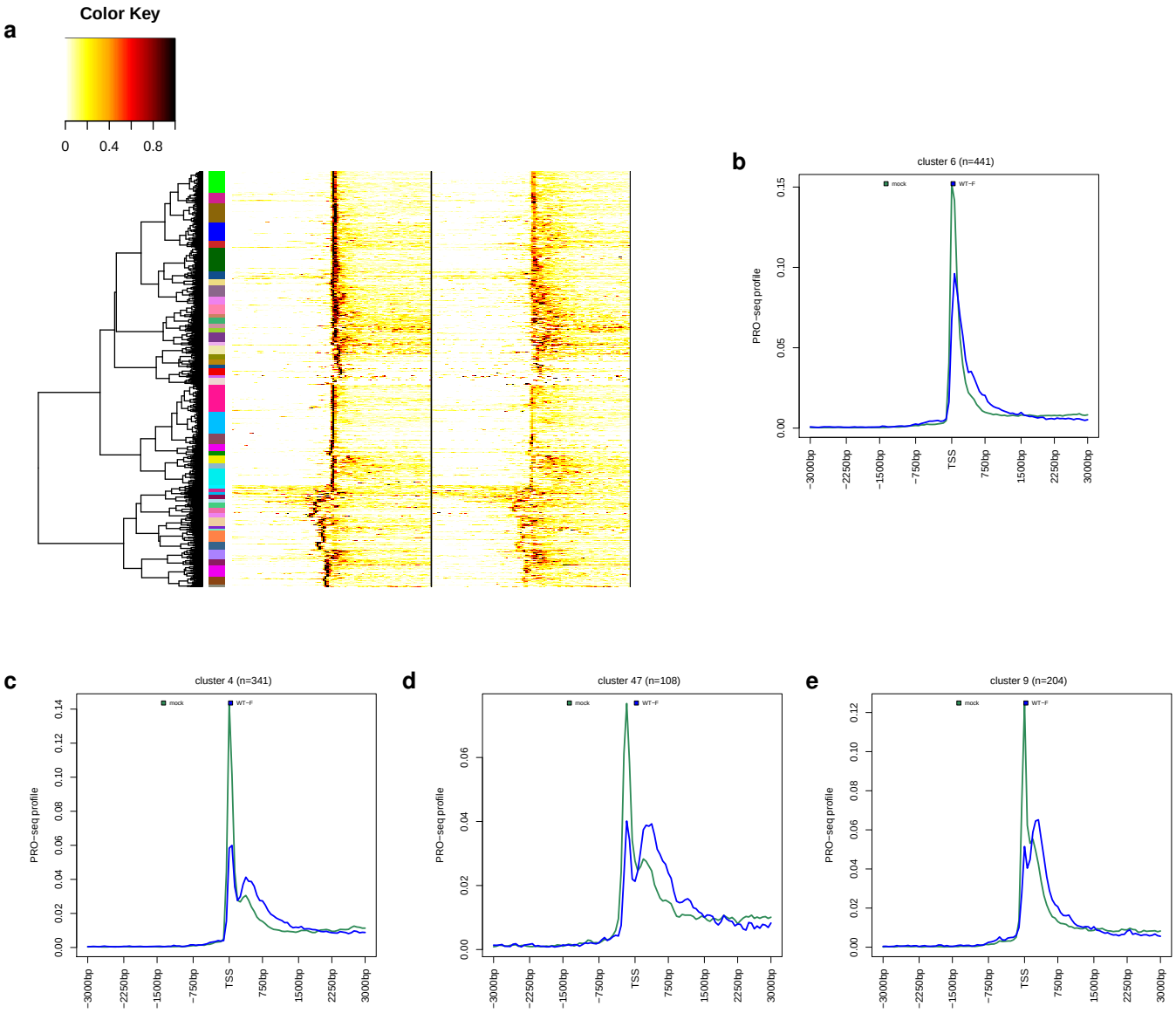

a
8192.00
256.00
8.00
0.25
WT−F 3 h p.i. (PI)
0.125
4.000
128.000
4096.000
mock infection (PI)
gene group
increased PI
slightly reduced PI
strongly reduced PI
b
4096.000
128.000
4.000
0.125
PI
mock
WT−F 3 h p.i.
gene group
increased PI
slightly reduced PI
strongly reduced PI
c
32768.000
512.000
8.000
0.125
gene body RPKM
mock
WT−F 3 h p.i.
gene group
increased PI
slightly reduced PI
strongly reduced PI
d
2
1
0
−1
−2
WT−17 vs. mock (log2 fold−change)
0−1 h p.i.
1−2 h p.i.
2−3 h p.i.
3−4 h p.i.
4−5 h p.i.
5−6 h p.i.
6−7 h p.i.
7−8 h p.i.
gene group
increased PI
slightly reduced PI
strongly reduced PI

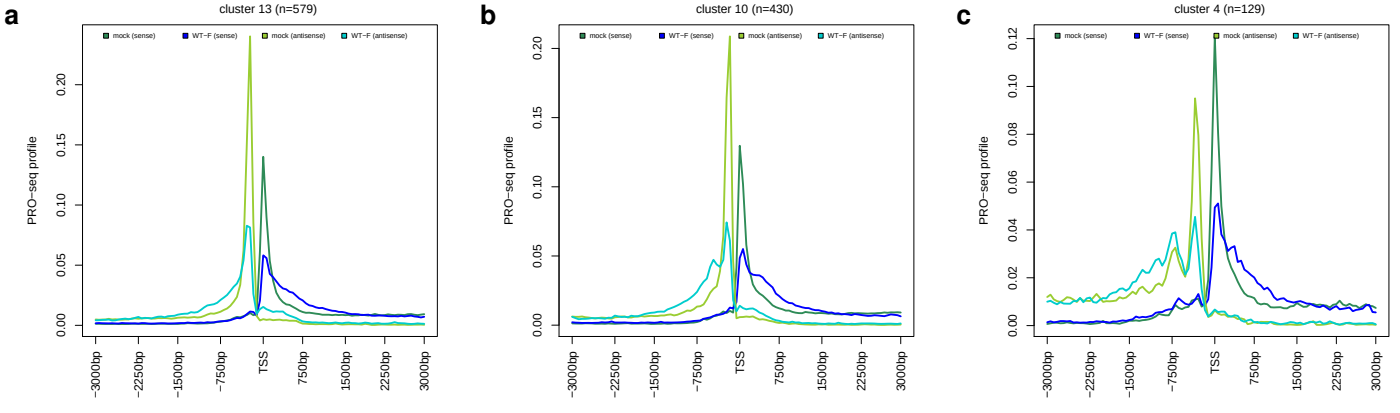

a
ATP5G1
ATP5G1
RP11-501C14.9
PRO-seq
mock [+]
[0, 5.4]
TSS
WT-F [+]
[0, 1.31]
TSS
46.967 mb
46.969 mb
46.971 mb
46.973 mb
46.975 mb
46.968 mb
46.97 mb
46.972 mb
46.974 mb
46.976 mb
METTL13
METTL13
PRO-seq
mock [+]
[0, 1.88]
TSS
WT-F [+]
[0, 0.63]
TSS
171.748 mb
171.75 mb
171.752 mb
171.754 mb
171.756 mb
171.758 mb
171.749 mb
171.751 mb
171.753 mb
171.755 mb
171.757 mb

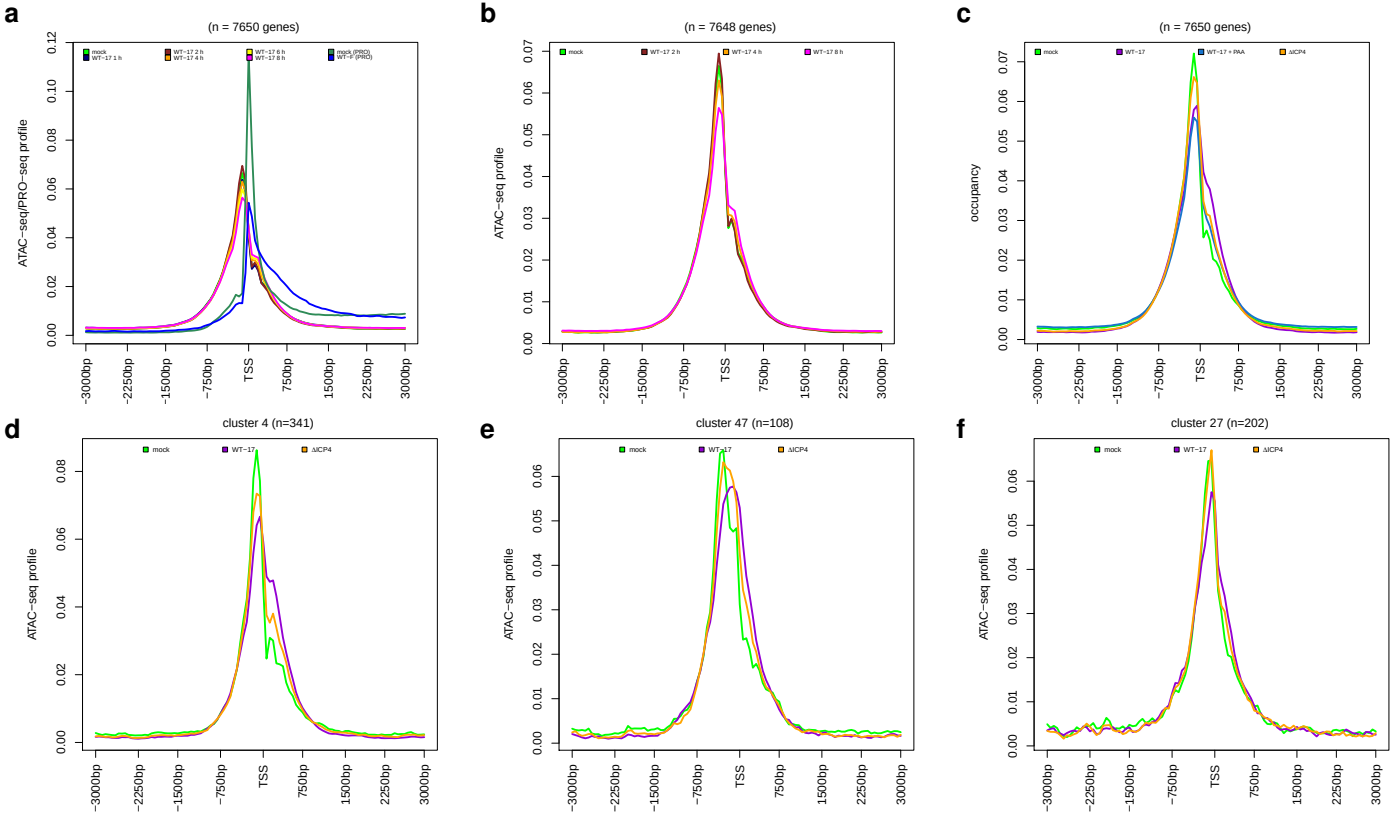

b
increased PI (n=763)
mock (sense)
WT−F (sense)
mock (antisense)
WT−F (antisense)
PRO−seq profile
0.00
0.01
0.02
0.03
0.04
0.05
0.06
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp

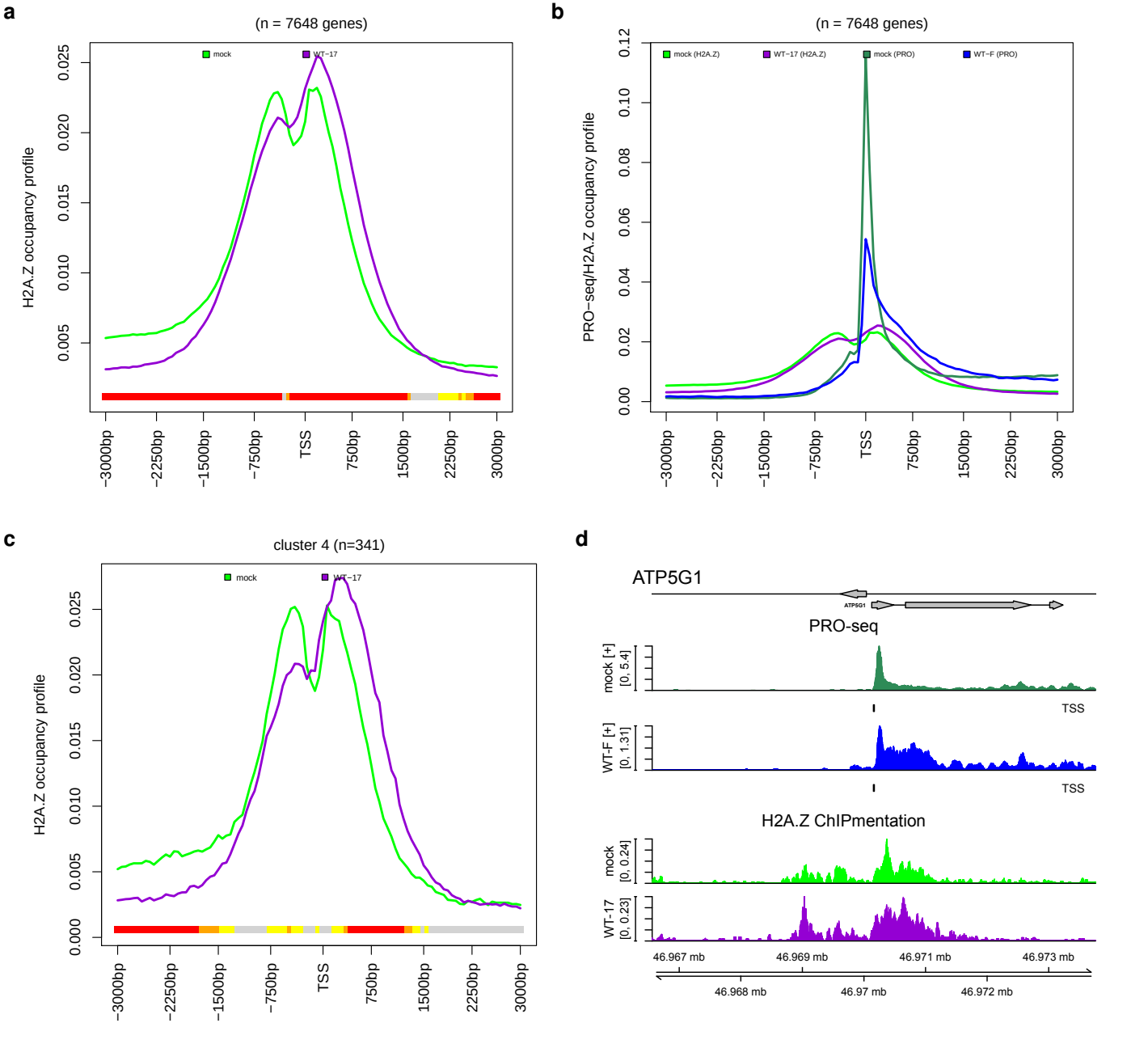

c
strongly reduced PI (n=4211)
mock (sense)
WT−F (sense)
mock (antisense)
WT−F (antisense)
PRO−seq profile
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp


d
slightly reduced PI (n=2082)
mock (sense)
WT−F (sense)
mock (antisense)
WT−F (antisense)
PRO−seq profile
0.00
0.02
0.04
0.06
0.08
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp


a
Color Key
0
0.4
0.8
b
cluster 6 (n=441)
mock
WT−F
PRO−seq profile
0.00
0.05
0.10
0.15
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
c
cluster 4 (n=341)
mock
WT−F
PRO−seq profile
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
d
cluster 47 (n=108)
mock
WT−F
PRO−seq profile
0.00
0.02
0.04
0.06
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
e
cluster 9 (n=204)
mock
WT−F
PRO−seq profile
0.00
0.02
0.04
0.06
0.08
0.10
0.12
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp


a
cluster 13 (n=579)
b
cluster 10 (n=430)
c
cluster 4 (n=129)
mock (sense)
WT−F (sense)
mock (antisense)
WT−F (antisense)
PRO−seq profile
0.00
0.05
0.10
0.15
0.20
0.02
0.04
0.06
0.08
0.12
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp


a
(n = 7650 genes)
mock
WT−17 1 h
WT−17 2 h
WT−17 4 h
WT−17 6 h
WT−17 8 h
mock (PRO)
WT−17 (PRO)
ATAC−seq/PRO−seq profile
0.00
0.02
0.04
0.06
0.08
0.10
0.12
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
b
(n = 7648 genes)
mock
WT−17 2 h
WT−17 4 h
WT−17 8 h
ATAC−seq profile
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
c
(n = 7650 genes)
mock
WT−17
WT−17 + PAA
ΔICP4
occupancy
d
cluster 4 (n=341)
mock
WT−17
ΔICP4
e
cluster 47 (n=108)
f
cluster 27 (n=202)


a
(n = 7648 genes)
mock
WT−17
H2A.Z occupancy profile
0.005
0.010
0.015
0.020
0.025
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
b
(n = 7648 genes)
mock (H2A.Z)
WT−17 (H2A.Z)
mock (PRO)
WT−F (PRO)
PRO−seq/H2A.Z occupancy profile
0.00
0.02
0.04
0.06
0.08
0.10
0.12
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
c
cluster 4 (n=341)
mock
WT−17
H2A.Z occupancy profile
0.000
0.005
0.010
0.015
0.020
0.025
−3000bp
−2250bp
−1500bp
−750bp
TSS
750bp
1500bp
2250bp
3000bp
d
ATP5G1
ATP5G1
PRO-seq
mock [+]
[0, 5.4]
TSS
WT-F [+]
[0, 1.31]
TSS
H2A.Z ChIPmentation
mock
[0, 0.24]
WT-17
[0, 0.23]
46.967 mb
46.969 mb
46.971 mb
46.973 mb
46.968 mb
46.97 mb
46.972 mb