

Differences in RNA polymerase II complexes and their interactions with surrounding chromatin on human and cytomegalovirus genomes

Benjamin M. Spector^a, Mrutyunjaya Parida^a, Ming Li^{a,b,c,d}, Christopher B. Ball^a, Jeffery L. Meier^{b,c,d}, Donal S. Luse^e, and David H. Price^a

^aDepartment of Biochemistry and Molecular Biology, The University of Iowa, Iowa City, IA 52242, USA

^bDepartment of Internal Medicine, The University of Iowa, Iowa City, IA 52242, USA

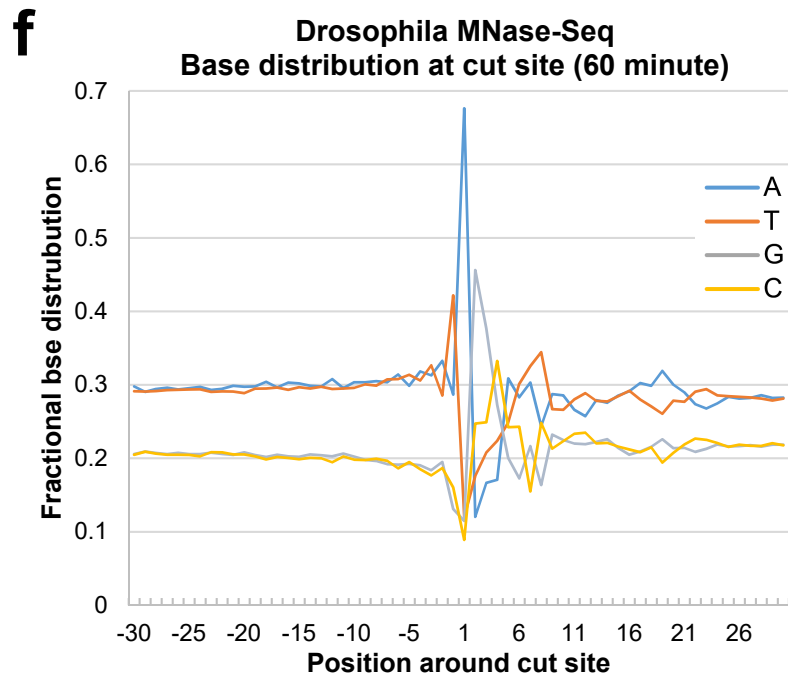
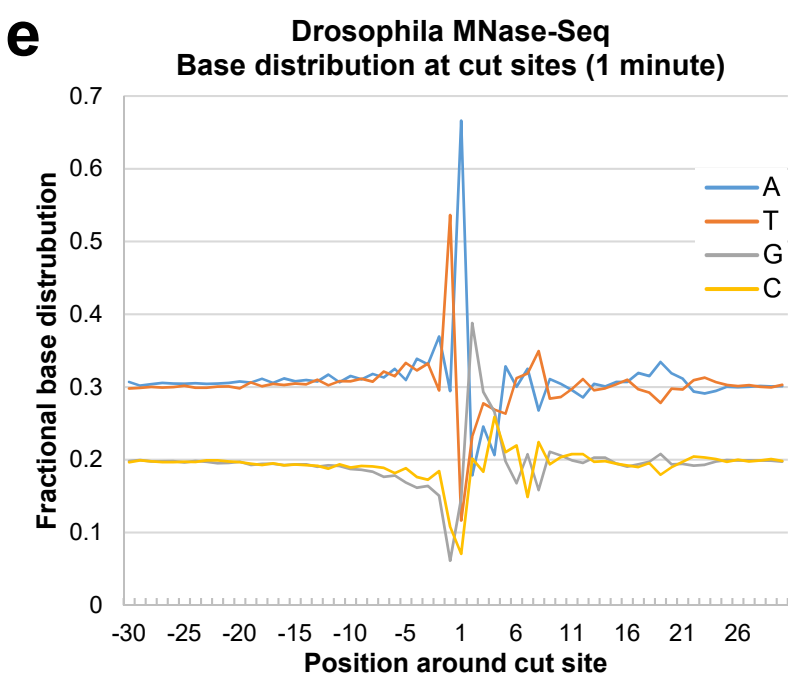
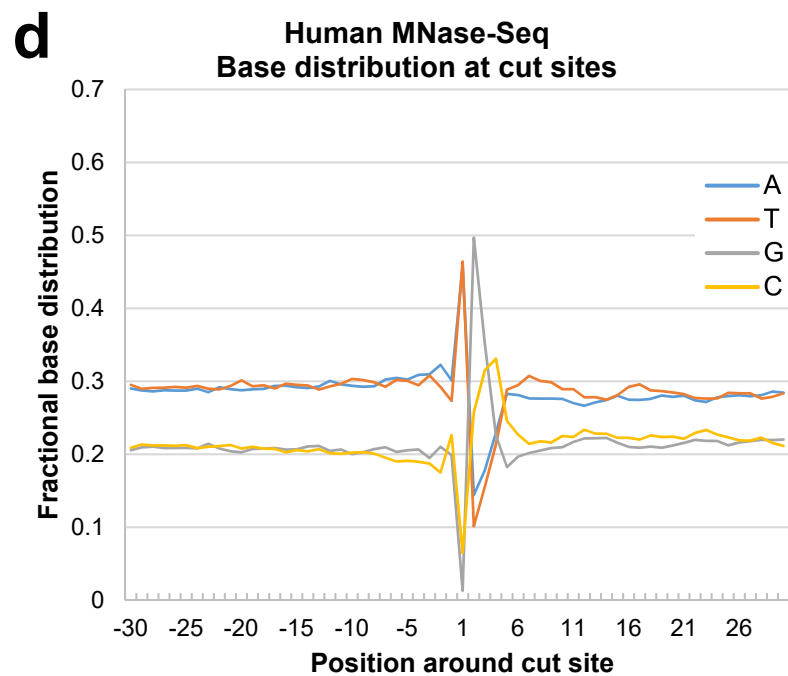
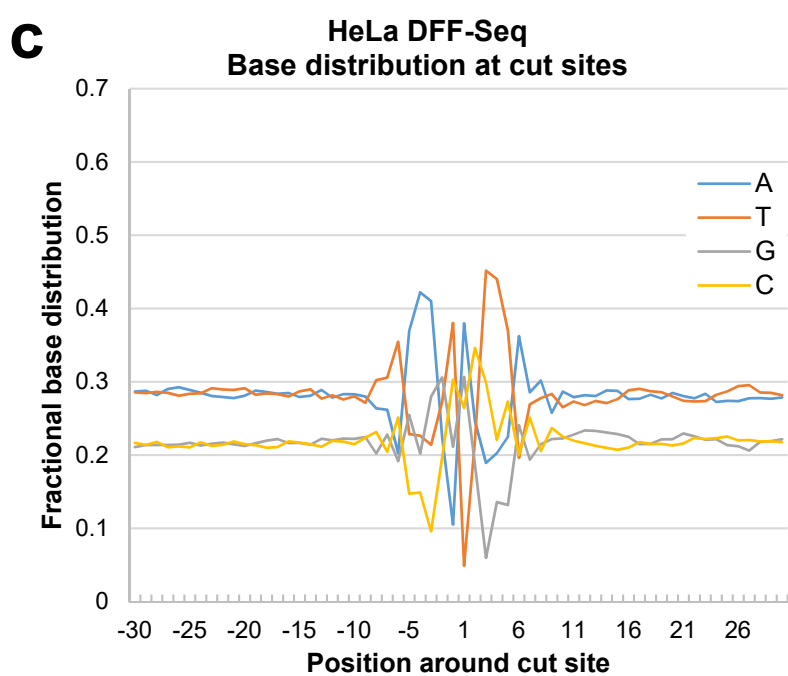
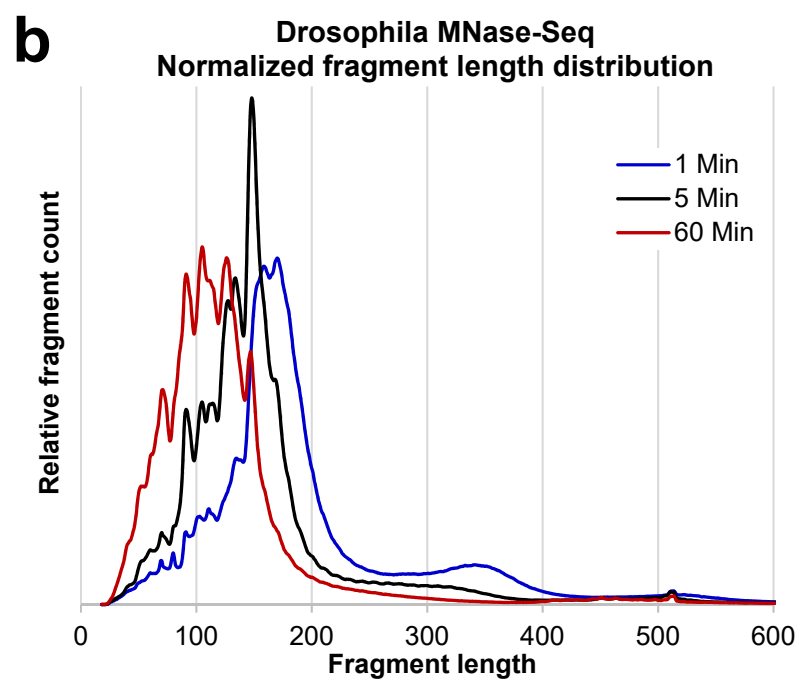
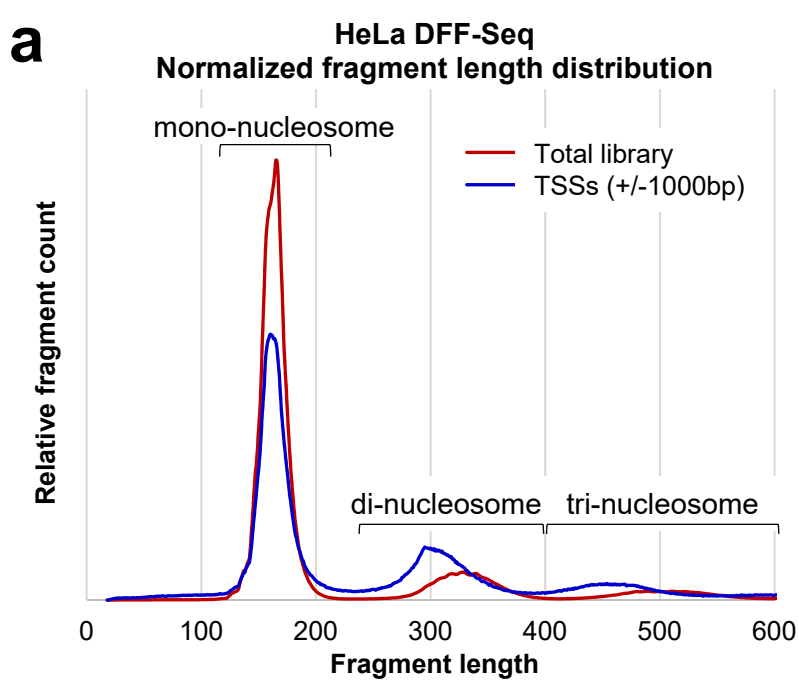
^cDepartment of Epidemiology, The University of Iowa, Iowa City, IA 52242, USA

^dVeterans Affairs Health Care System, Iowa City, IA 52242, USA

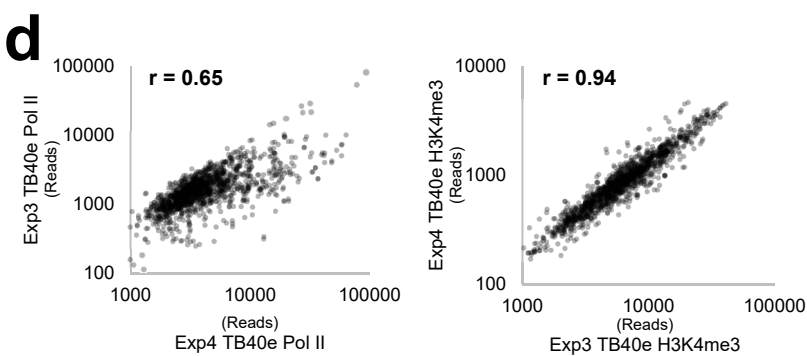
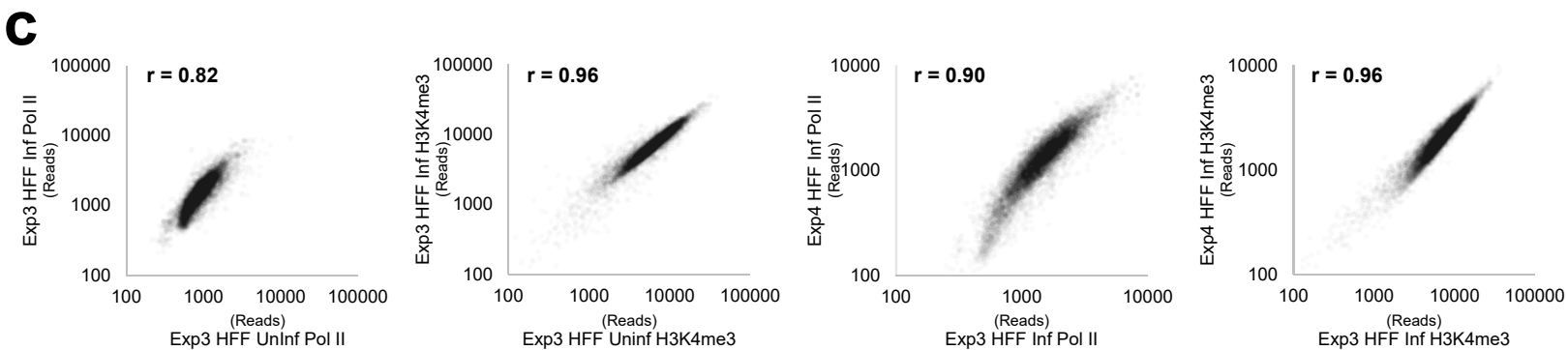
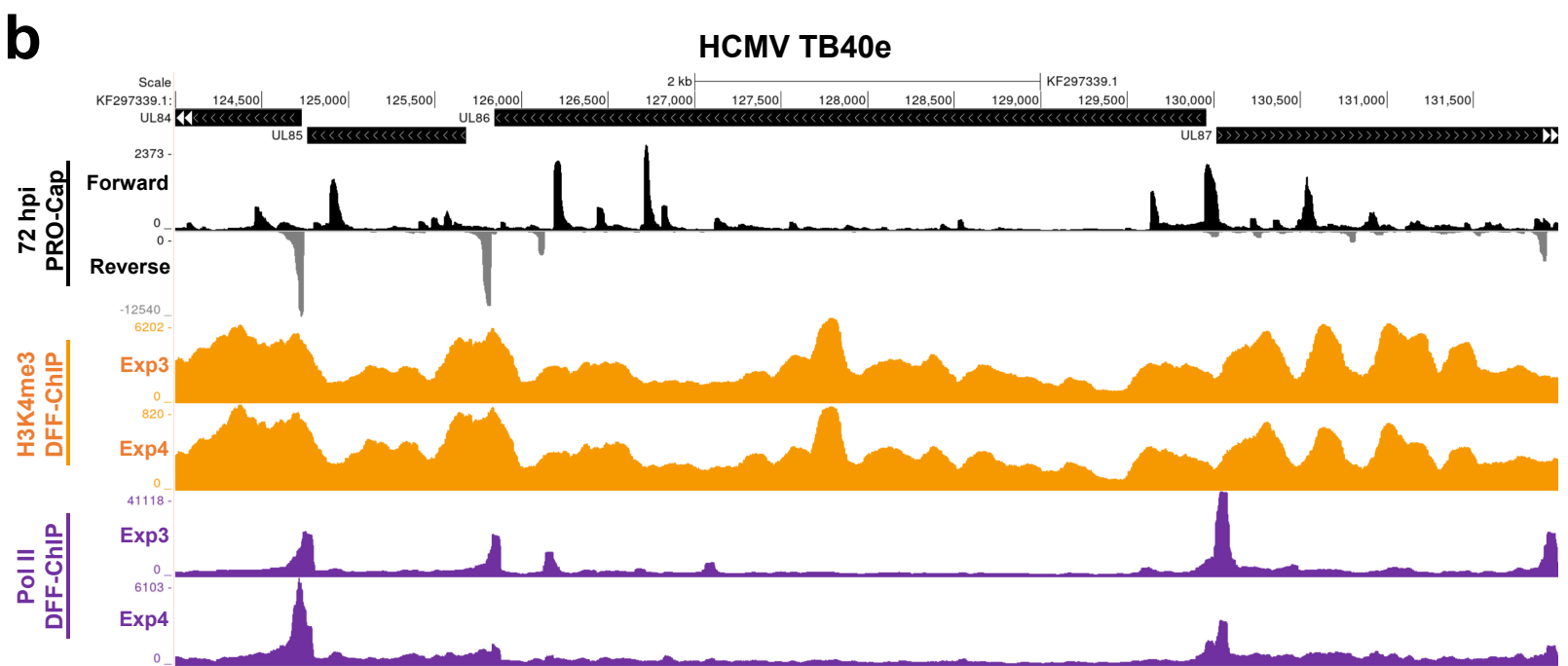
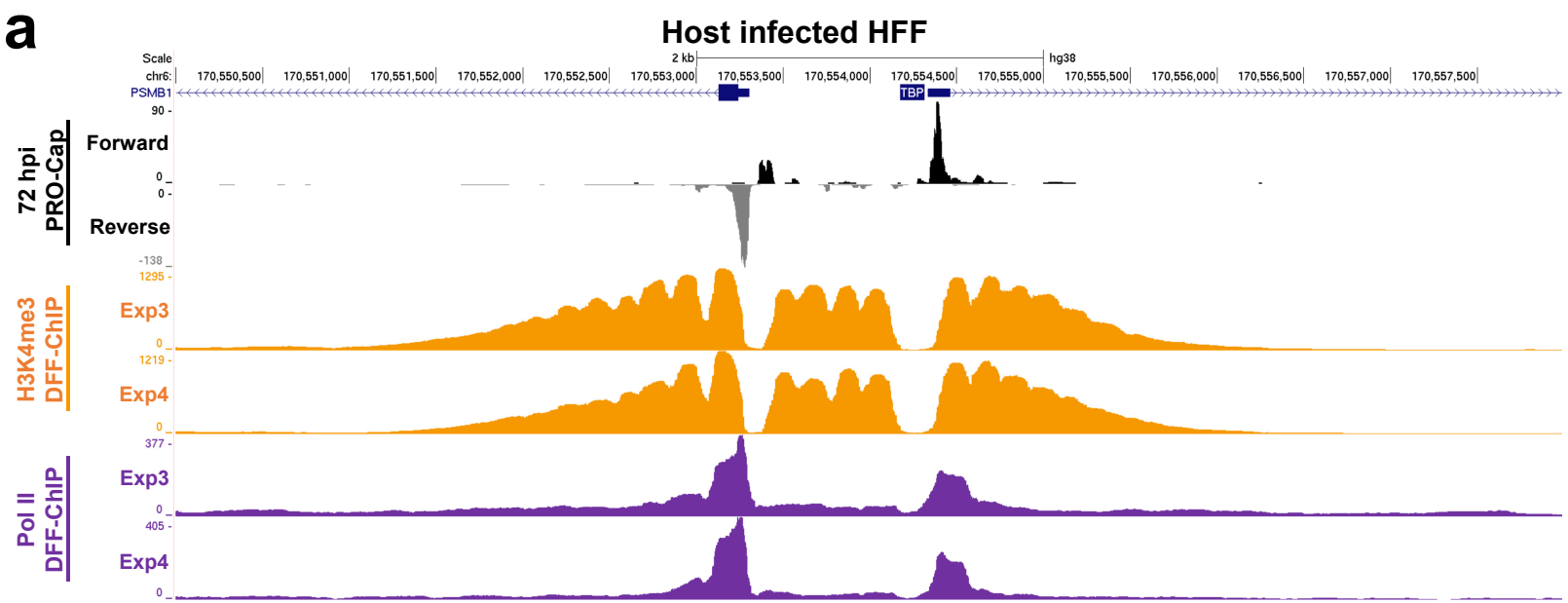
^eDepartment of Cardiovascular and Metabolic Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, 44195, USA.

Keywords: RNA polymerase II, preinitiation complexes, chromatin, human cytomegalovirus, TBP, UL87, DFF ChIP-Seq

Supplemental Figures

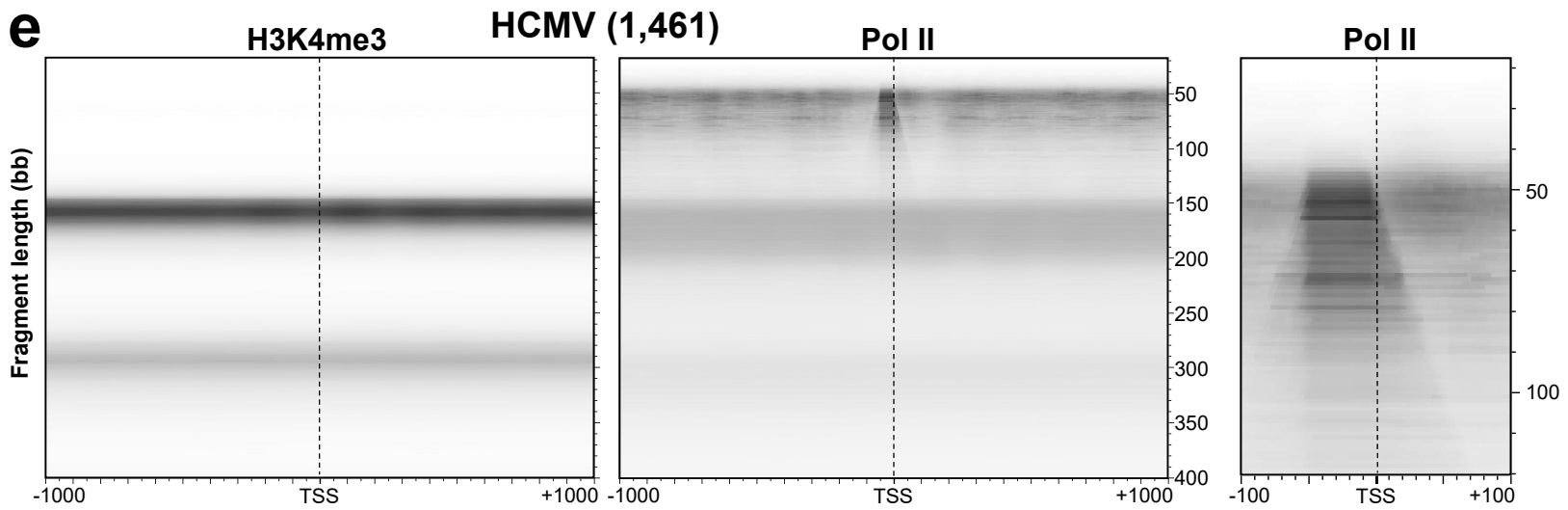
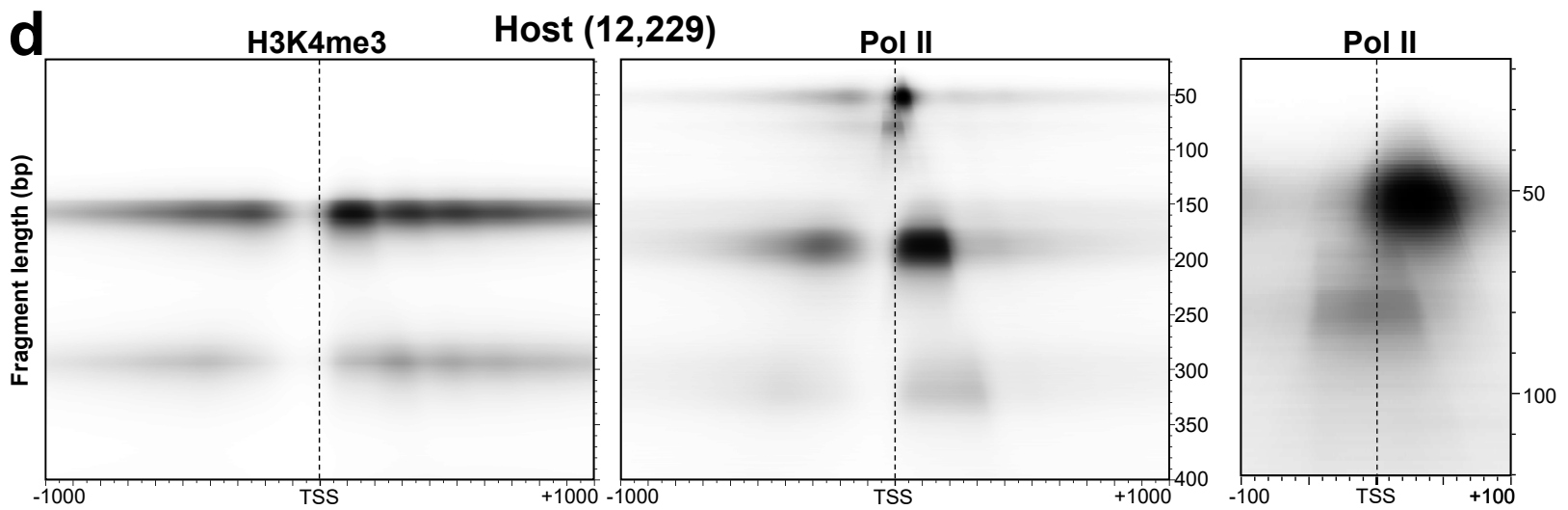
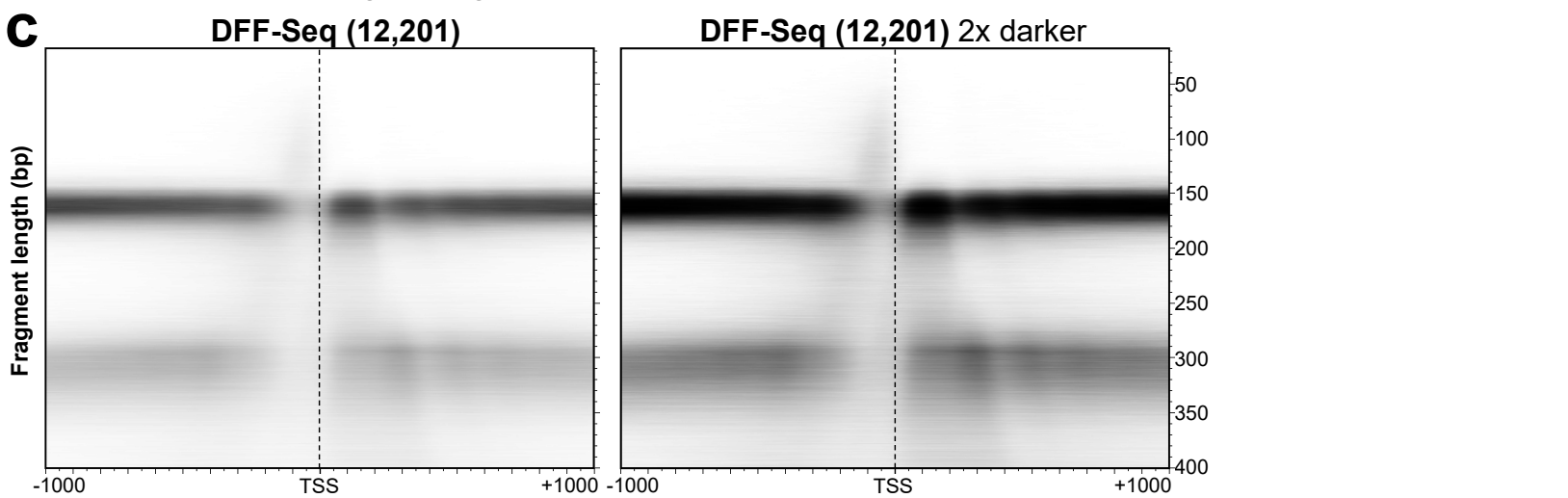
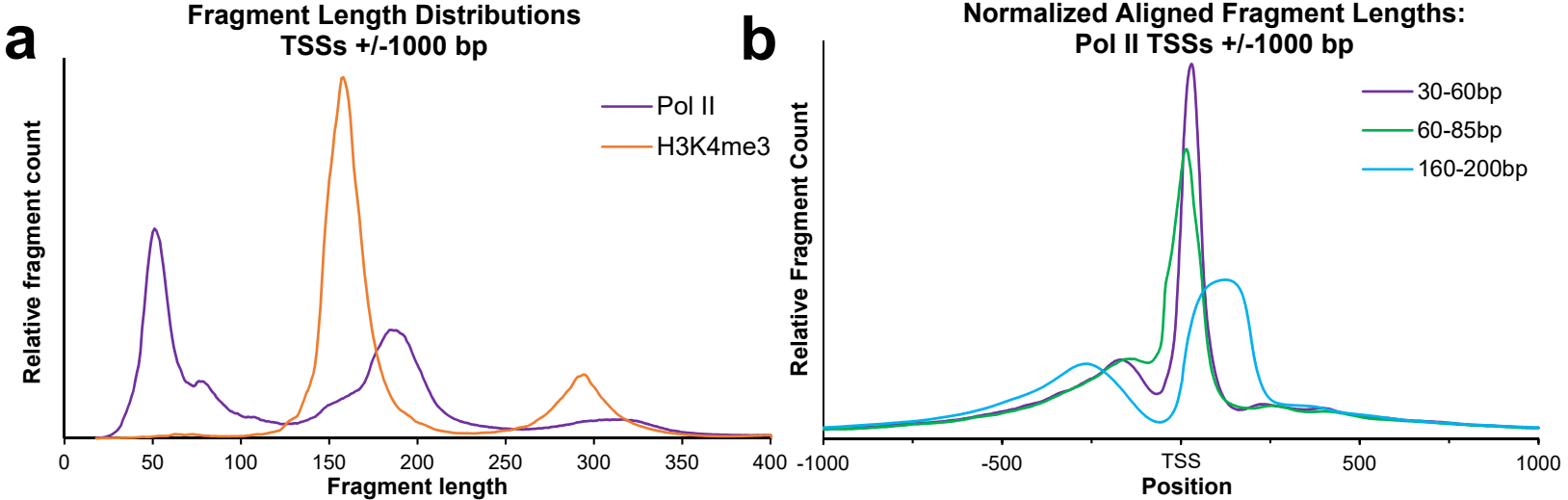


Supplementary Fig. 1. Analysis of DFF-Seq and comparison of DFF-Seq to MNase-Seq
a Fragment length distribution for the total DFF-Seq library⁶ and for fragments +/- 1000 bp from the MaxTSS of 12,201 truQuant HeLa promoters. Fragment counts were normalized. Mono-, di-, and tri-nucleosome are demarcated. **b** Fragment length distribution of MNase-Seq data generated in *Drosophila* utilizing 1 min, 5 min, and 60 min time points of digestion⁵¹. The three curves were normalized by making the area under curves equal. **c** Base distribution of the surrounding 60 bp around the 5' ends DFF-Seq fragments. **d** Base distribution of the surrounding 60 bp around the 5' ends of MNase-Seq fragments from human cells³⁰. **e,f** Base distribution of the surrounding 60 bp around the 5' ends of MNase-Seq fragments generated from the 1 minute and 60-minute time point of digestion in *Drosophila*⁵¹.



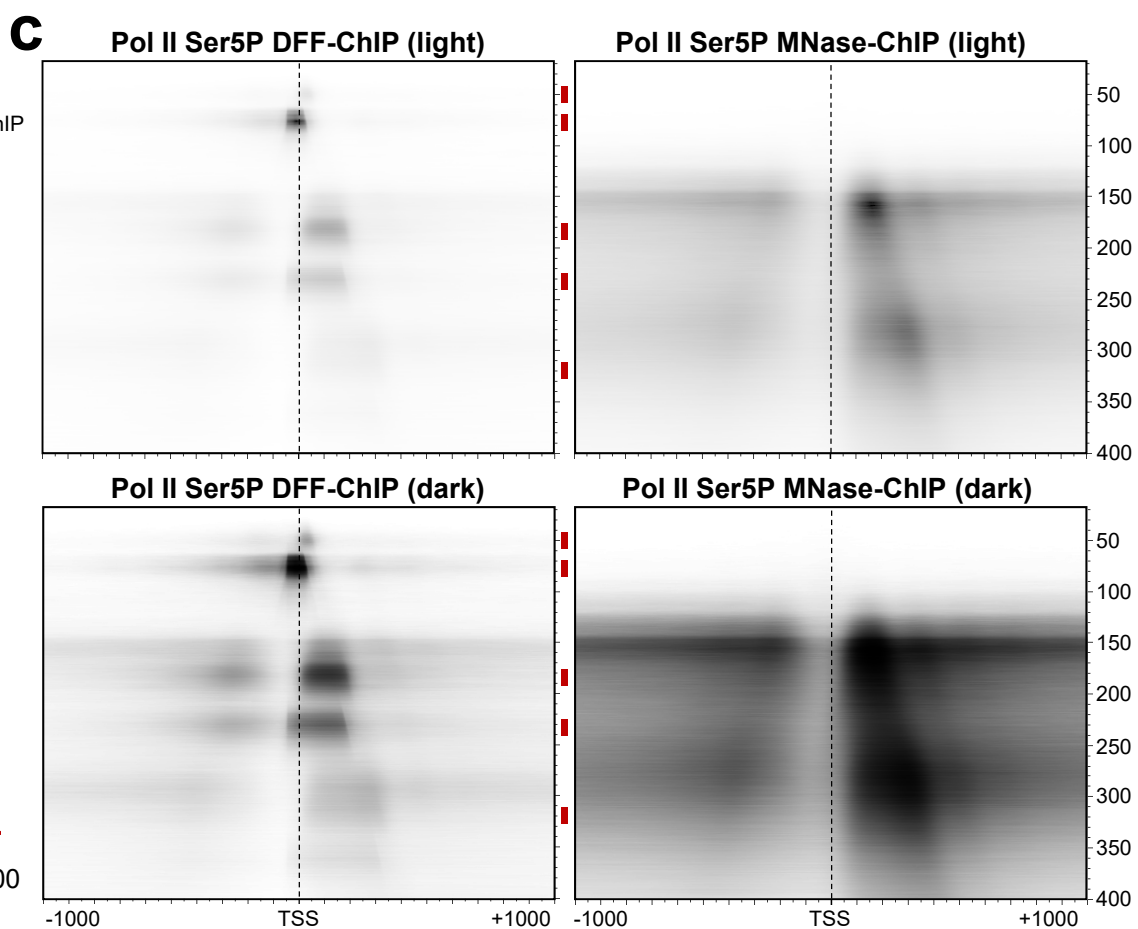
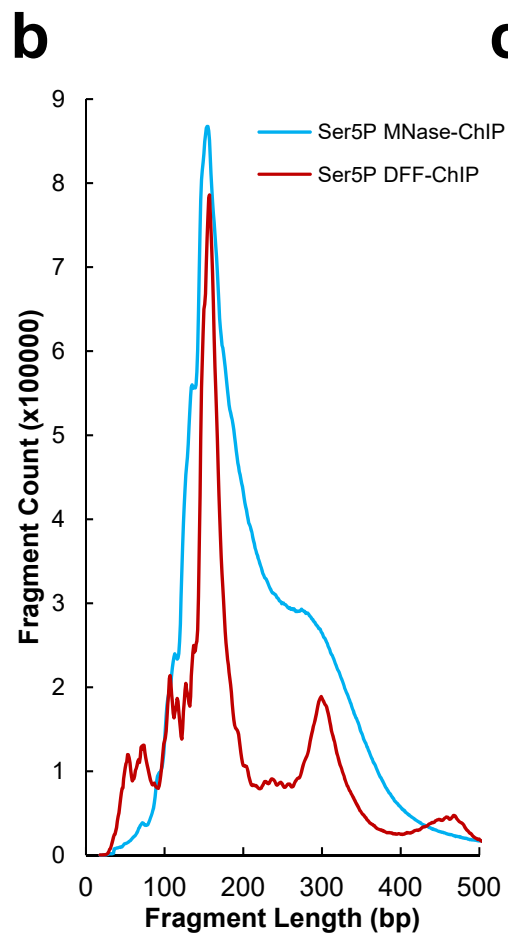
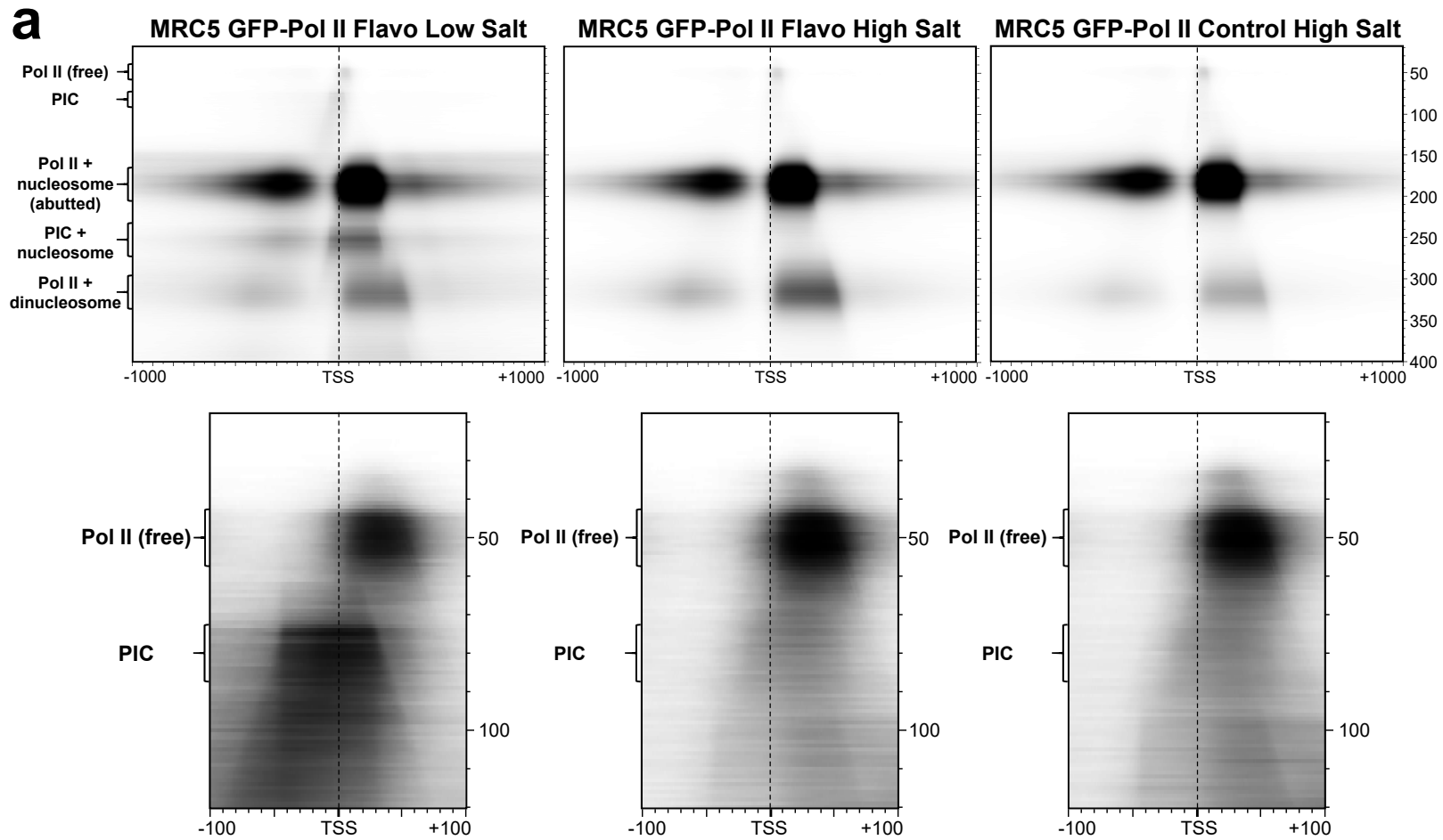
Supplementary Fig. 2. Reproducibility of DFF-ChIP on infected HFFs

a, b Genome browser tracks DFF-ChIP H3K4me3 and Pol II tracks from Exp3 and Exp4 and 72 hpi PRO-Cap transcription data. Representative 8000 bp regions are shown on the host and HCMV genomes. **c,d** Correlations between Exp3 and Exp4 datasets. Host reads in 10,000 bp windows centered on the MaxTSS of 12,229 HFF truQuant were summed and plotted against sums from other datasets (**c**). Viral reads from 1,461 200 bp transcription start regions (TSRs)²⁵ identified using TSR-finder were summed and plotted against sums from other datasets (**d**). Pearson correlations (*r*) are provided.



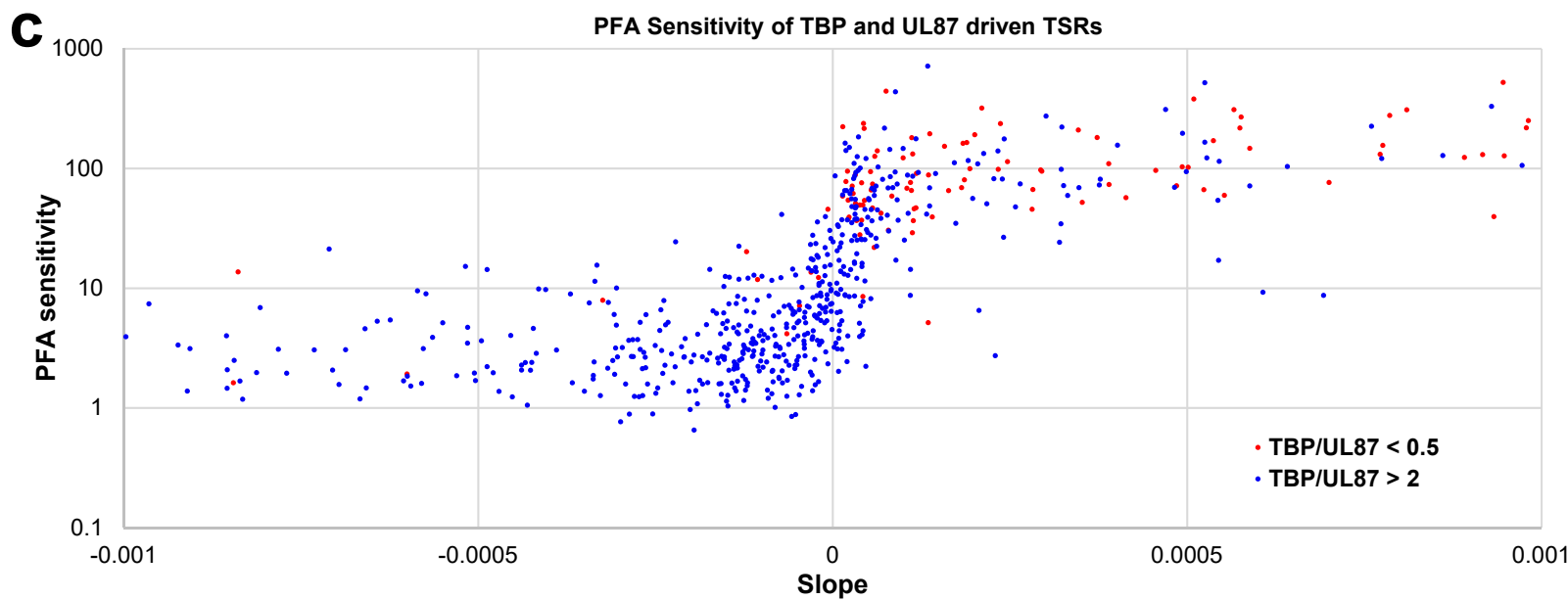
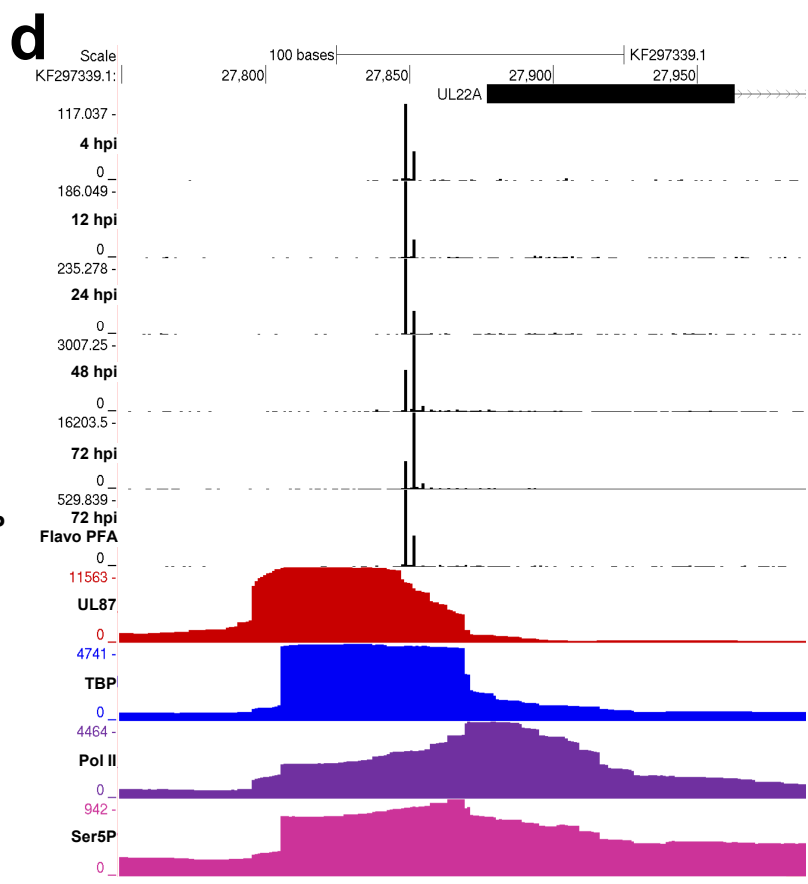
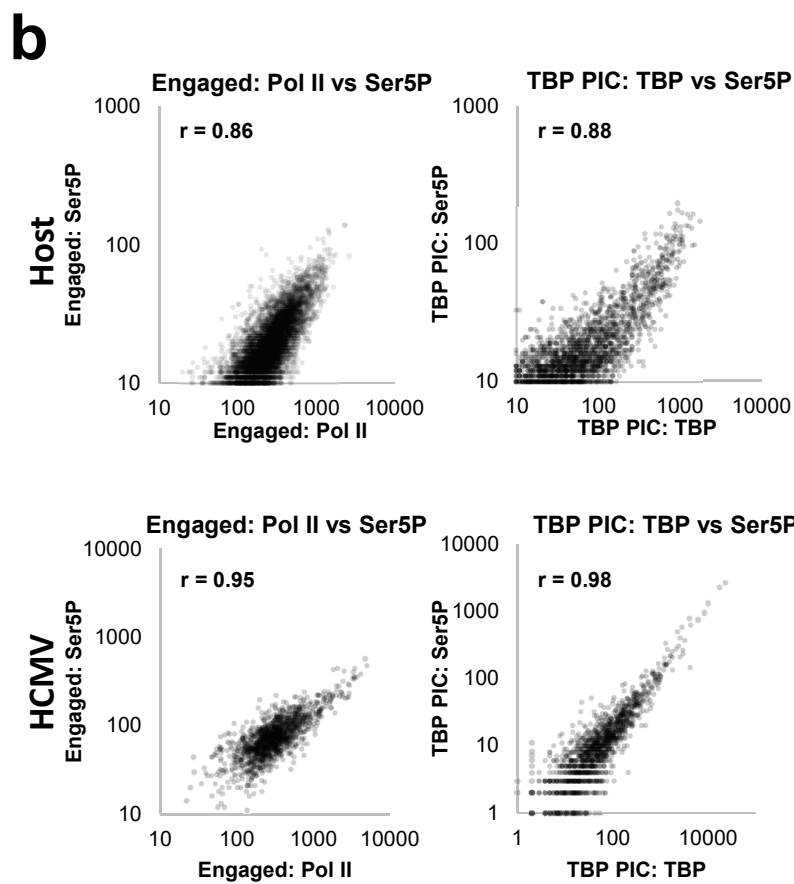
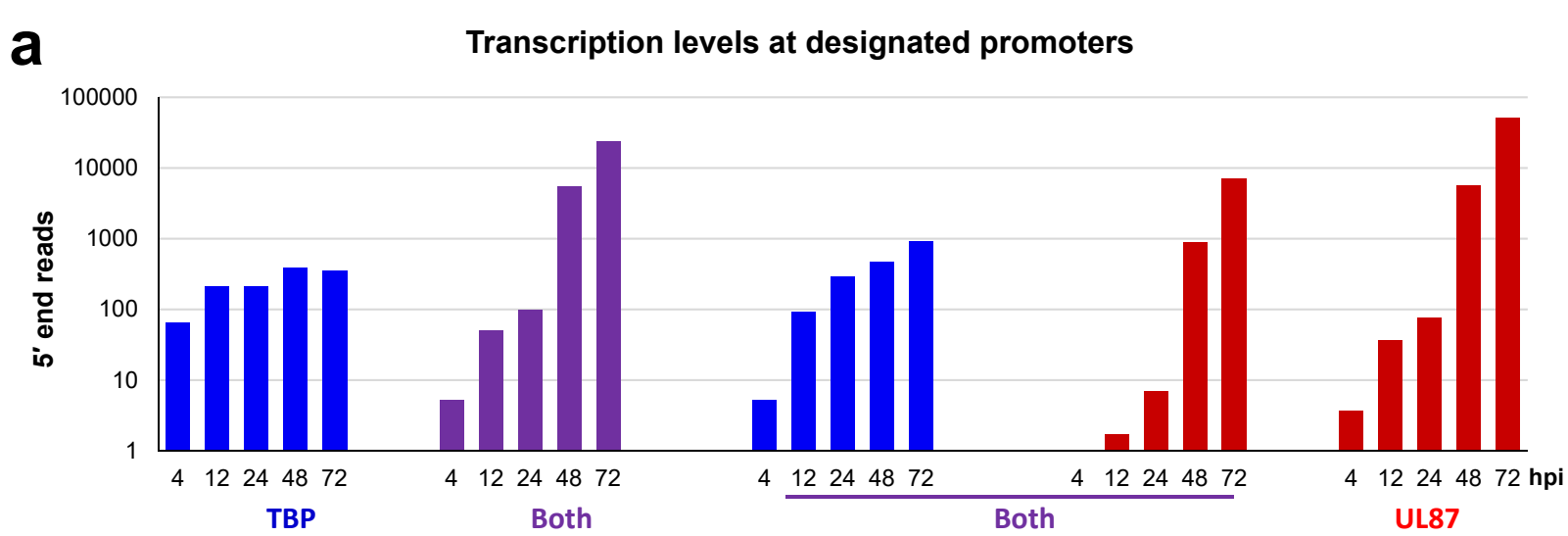
Supplementary Fig. 3. Visualizing transcription complexes and chromatin from DFF-Seq and Exp3 replicate data

a Length distribution of fragments +/- 1000 bp of each of 12,229 truQuant promoters for the Pol II and H3K4me3 DFF-ChIP datasets (Exp3). **b** Fragment count of selected ranges of fragments from the Pol II DFF-ChIP (Exp3) positioned across the 2000 bp region centered on the MaxTSS of 12,229 HFF truQuant promoters. The three curves were normalized by making the area under the curves equal. **c** fragMap representation of DFF-Seq data shown using two different black values for the 12,201 HeLa truQuant promoters showing fragments from 18-400 bp that are positioned +/-1000 bp around the MaxTSS. **d** FragMaps of H3K4me3 (Exp3) and Pol II (Exp3) for the 12,229 HFF truQuant promoters showing 18-400 bp fragments that are +/-1000 bp around the MaxTSS. A zoomed fragMap showing 18-120 bp fragments positioned +/-100 bp around the MaxTSS is also shown (right). **e** FragMaps of H3K4me3 (Exp3) and Pol II (Exp3) for the 1,461 TB40/E TSRs showing 18-400 bp fragments positioned +/-1000 bp around the MaxTSS. A zoomed fragMap showing 18-120 bp fragments positioned +/-100 bp around the TSS is also shown (right).



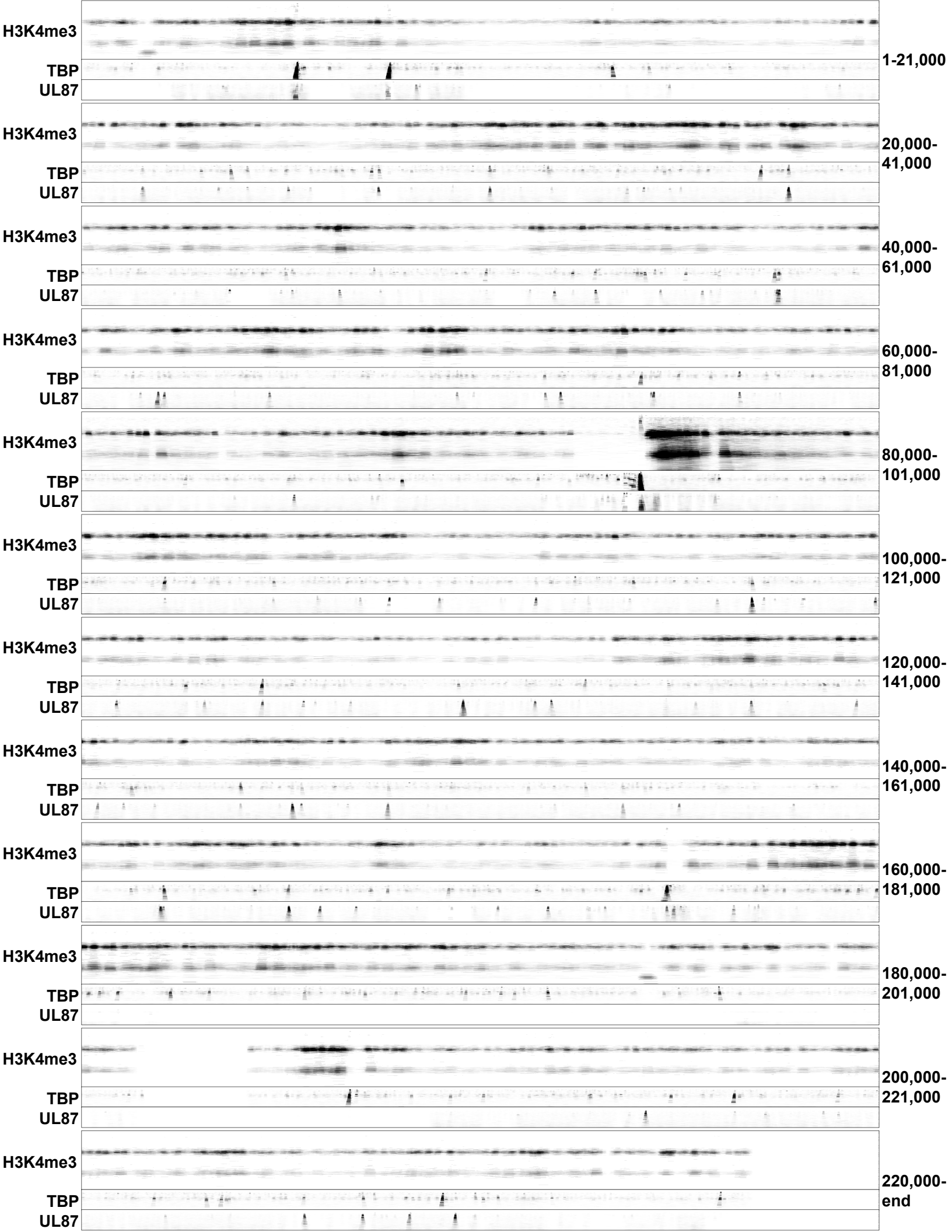
Supplementary Fig. 4. Salt sensitivity of TBP PICs on the host genome and inability of MNase-ChIP to detect transcription complexes.

a FragMaps of MRC5 GFP-Pol II datasets from Exp1 (flavo, low salt IP), Exp2 (flavo, high salt IP), and Exp2 (control high salt IP), showing 18-400 bp fragments positioned +/-1000 bp around the MaxTSS and zoomed in fragMaps showing 18-120 bp fragments positioned +/- 100 bp around the MaxTSS from 12,201 HeLa truQuant promoters. Fragment lengths of specified features are demarcated by brackets. For the wider fragMaps, the black value was reduced 3 fold to emphasize smaller fragments. For the zoomed in fragMaps, max black levels were set based on the max value in the shown region as standard for fragMaps. **b** Fragment distributions of the entire Exp4 Ser5P DFF-ChIP (red) or Ser5P MNase-ChIP libraries (blue)⁵². **c** FragMaps were created for Ser5P DFF-ChIP and Ser5P MNase-ChIP datasets and displayed with normal maximum black values (top) as well as with lower maximum black values to better visualize weaker features (bottom). Fragment sizes for transcription complexes (labeled in **a**) that were missing in the MNase-ChIP fragMap as compared to the DFF-ChIP fragMap are marked by red boxes.



Supplementary Fig. 5. TBP PICs are functionally different from UL87 PICs even though both may appear at late promoters

a Quantification of 5' end reads (TSSs) in an 11 bp window around the five MaxTSSs specifically indicated in Figure 5A. **b** Correlations of the amount of engaged Pol II and TBP PIC features counted utilizing the Ser5P dataset in comparison to the Pol II and TBP dataset, respectively, on the host and HCMV genomes (Exp3). Pearson correlations (r) are provided. **c** TSRs with a TBP/UL87 ratio greater than 2 or less than 0.5 from Fig. 5B were plotted using their PFA sensitivity against the value of their slope calculated from the read counts from the 5 PRO-Seq time points. TSRs with a TBP/UL87 ratio greater than 2 were colored blue and TSRs with a ratio less than 0.5 were colored red. **d** Genome browser tracks depicting shifting 5' ends for the UL22A promoter from PRO-Seq HCMV infection time course data and DFF-ChIP tracks showing UL87, TBP, Pol II, and Ser5P datasets (Exp4).

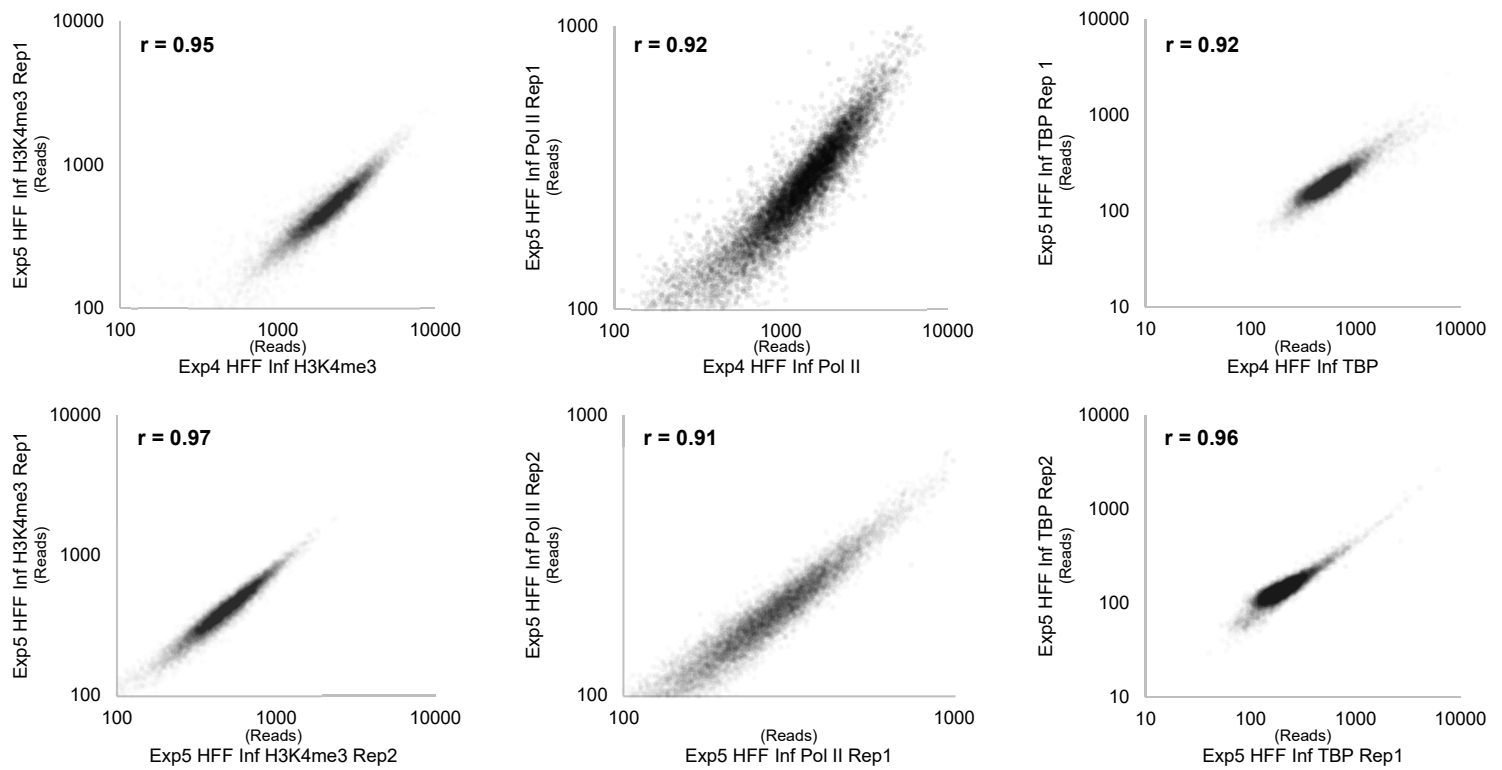


Supplementary Fig. 6. Genomic fragMaps spanning the entire HCMV genome for H3K4me3, TBP, and UL87 fragments

Genomic fragMaps depicting overlapping 21,000 bp windows of the entire HCMV genome. H3K4me3 depict fragments between 18-400 bp whereas TBP and UL87 fragMaps depict fragments between 18-150 bp. Note that there is a 1000 bp overlap between each region shown to allow better visualization of features at the beginning of each region.

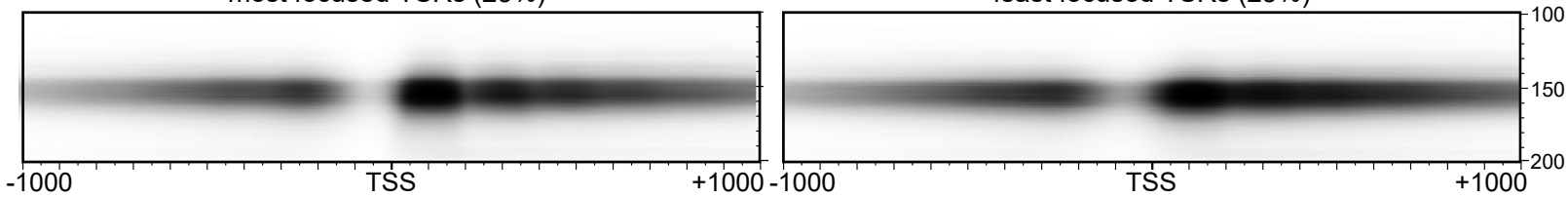
a

Correlations of host reads for 10 kb window around the MaxTSS of truQuant promoters

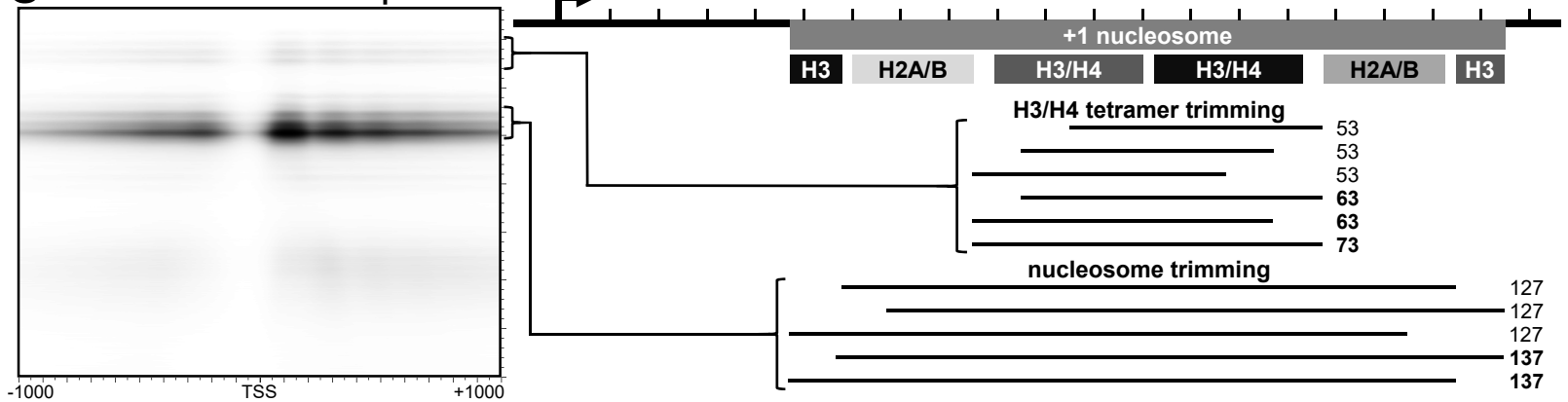
**b**

H3K4me3 host Exp4
most focused TSRs (25%)

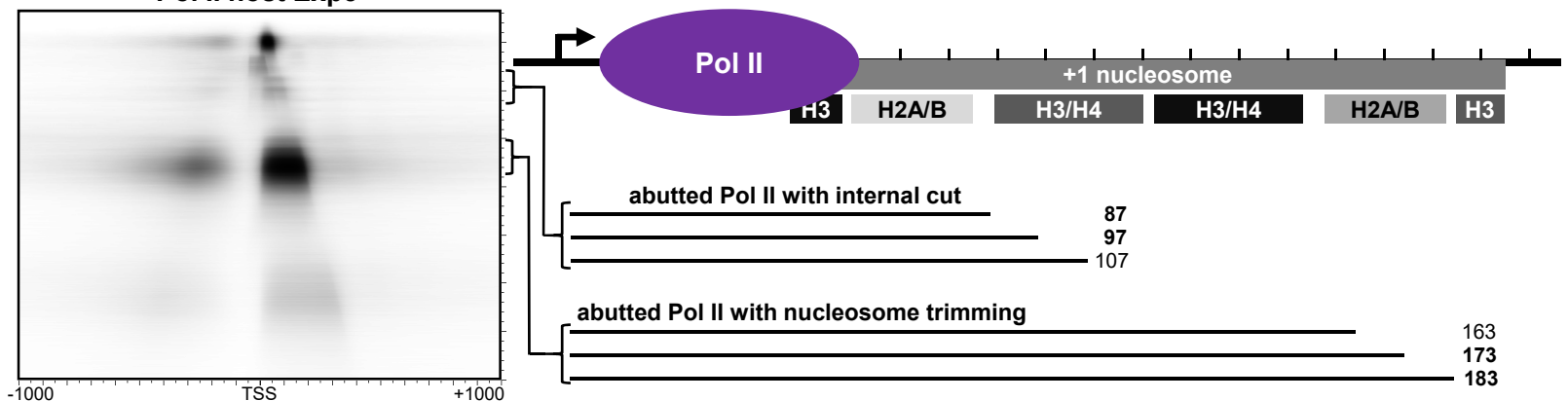
H3K4me3 host Exp4
least focused TSRs (25%)

**c**

H3K4me3 host Exp5



Pol II host Exp5



Supplementary Fig. 7. Characterization of fragments resulting from excess DFF digestion

a Correlation plots comparing number of reads in 10,000 bp windows centered on the MaxTSS of 12,229 HFF truQuant promoters from the indicated Exp4 and Exp5 datasets. Pearson correlations (r) are provided. **b** H3K4me3 (Exp4) fragMaps from the top 25% and bottom 25% of the 12,229 truQuant promoters sorted by TSS focus (standard deviation of TSSs in TSRs). **c** fragMaps of H3K4me3 (Exp5) and Pol II (Exp5) showing 18-400 bp fragments positioned +/- 1000 bp around the MaxTSS containing from 12,229 HFF truQuant promoters with a schematic depicting the most likely nucleosome protections and DFF cleavage sites. Exp5 datasets are each a combination of two replicas ($n = 2$).