Herpes Simplex viral nucleoprotein creates a competitive transcriptional environment facilitating robust viral transcription and host shut off

Sarah E. Dremel and Neal A. DeLuca*

Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA, 15219

Short title: HSV-1 ICP4 as nucleoprotein and transcription factor.

Key words: HSV1, ICP4, Pol II, transcription, genome replication

* Corresponding author:

Neal A. DeLuca

Department of Microbiology and Molecular Genetics

University of Pittsburgh School of Medicine

514 Bridgeside Point II

450 Technology Dr.

Pittsburgh, PA 15219

E-mail: ndeluca@pitt.edu

Phone:(412) 648-9947

FAX: (421) 624-1401

Highlights (85 characters per highlight)

- HSV-1 ICP4 coats the viral genome promoting robust recruitment of Pol II transcription machinery.
- ICP4 prefers the viral genome due to the absence of nucleosomes and density of binding motifs.
- At high concentrations ICP4 promiscuously binds DNA including euchromatic host promoters.
- ICP4 is required for host transcriptional shut off, independent of genome replication.

Summary (≤150 words)

Herpes simplex virus-1 (HSV-1) replicates within the nucleus coopting the host's RNA Polymerase II (Pol II) machinery for production of viral mRNAs culminating in host transcriptional shut off. The mechanism behind this rapid reprogramming of the host transcriptional environment is largely unknown. We identified ICP4 as responsible for preferential recruitment of the Pol II machinery to the viral genome. ICP4 is a viral nucleoprotein which binds double stranded DNA. We determined ICP4 discriminately binds the viral genome due to the absence of cellular nucleosomes and high density of cognate binding sites. We posit that ICP4's ability to recruit not just Pol II, but also more limiting essential components, such as TBP and Mediator create a competitive transcriptional environment. These distinguishing characteristics ultimately result in a rapid and efficient reprogramming of the host's transcriptional machinery, which does not occur in the absence of ICP4.

INTRODUCTION

Like most DNA viruses, the genome of Herpes simplex virus-1 (HSV-1) is transcribed by RNA Polymerase II (Pol II) (Alwine et al., 1974). It's approximately 85 genes (McGeoch et al., 1988; McGeoch et al., 1986; McGeoch et al., 1985) are transcribed in a temporally coordinated sequence, such that their protein products are expressed at the appropriate time in the life cycle of the virus (Honess and Roizman, 1974a; Honess and Roizman, 1974b; Honess and Roizman, 1975). Immediate early (IE) gene products enable the efficient expression of early (E) and late (L) genes. The protein products of E genes are mostly involved in DNA replication. DNA replication and IE proteins enable the efficient transcription of L genes, which encode the structural components of the virus. DNA replication licenses L promoters, enabling the binding of core Pol II transcription factors, thus activating the initiation of L transcription (Dremel and DeLuca, 2019). This entire transcriptional cascade is observed within 3 hours (h) post entry (Dembowski and DeLuca, 2018b; Dremel and DeLuca, 2019), culminating in production of the first viral progeny between 4 and 6 h post-infection (hpi). To accomplish this robust and rapidly changing program of transcription, the viral genome must compete with the vastly larger cellular genome for numerous Pol II transcription factors, in addition to mediating the possible constraints of cellular histones.

A major component of this cascade is the IE protein Infected Cell Polypeptide 4 (ICP4) (Courtney and Benyesh-Melnick, 1974). ICP4 is essential for viral growth because it promotes efficient transcription of viral E and L genes (Dixon and Schaffer, 1980; Preston, 1979; Watson and Clements, 1980). Thus, in the absence of ICP4, E and L proteins are poorly expressed, IE proteins are overproduced, DNA replication does not occur, and there is no detectable viral yield (DeLuca et al., 1985). ICP4 was first shown to bind to DNA cellulose made from salmon sperm DNA (Powell and Purifoy, 1976). Faber and Wilcox later showed ICP4 has sequence-specific DNA binding activity (Faber and Wilcox, 1986). ICP4 interacts

with a number of cellular general transcription factors (GTFs), predominantly components of TFIID and the Mediator complex (Carrozza and DeLuca, 1996; Lester and Deluca; Wagner and DeLuca, 2013), facilitating their recruitment to the viral genome through its DNA binding activity (Dembowski and DeLuca, 2018a; Lester and Deluca; Sampath and Deluca, 2008). ICP4 is synthesized early in infection, binds to the viral genome located at ND10 structures (Everett et al., 2003), and remains associated with the genome throughout all phases of infection (Dembowski and DeLuca, 2018a). Therefore, ICP4 has the potential to influence events occurring on the viral genome from a time when genome number is at a minimum, and ICP4 expression is peaked, through a time when genome numbers are greatly elevated by replication.

Studies have also shown that epigenetic modulation of histones associated with the viral genome early in infection can affect productive viral infection (Knipe and Cliffe, 2008; Liang et al., 2009). However, we have shown that the abundance of histones is relatively low or absent, and that ICP4 is one of the most abundant proteins on viral genomes during productive infection (Dembowski and DeLuca, 2015, 2018a; Dembowski et al., 2017). In this study, we set out to determine the relationship between ICP4 and histones binding to the viral and cellular genomes, and the consequences for viral and cellular transcription. We propose that ICP4 is a major component of viral nucleoprotein, which functions in place of traditional cellular chromatin, and allows for the robust recruitment of cellular transcription factors specifically to the viral genome.

RESULTS

ICP4 binding is altered by viral genome replication.

Given the central role of ICP4 in viral gene transcription at all stages of infection, we were interested in how ICP4 interacts with the virus genome as the number exponentially increases, as a consequence of replication. We infected human fibroblast (MRC5) cells with wild-type HSV-1 (KOS) for 2, 4, and 6 hpi and performed ChIP-Seq for ICP4. Each time point represents a different replication state: 2 hpi (prereplication), 4 hpi (3-4 genome duplications), 6 hpi (5-6 genome duplications) (Fig. 1A). To quantitatively compare samples, we had to account for viral genome replication. Input samples provided the relative number of viral genomes present at each time point. We used this ratio to normalize immunoprecipitated (IP) sample for the amount of factor per genome. Early during infection (2h) ICP4 densely coated the viral genome (Fig. 1C). Viral genome replication decreased the amount of ICP4 bound per genome (Fig. 1A) resulting in a pattern containing sharper peaks. By 6 hpi ICP4 binding was retained exclusively on strong ICP4 binding motifs (Fig. 1C). Some of the retained binding sites were those previously established as having an inhibitory effect on the gene promoter bound, including ORF P, ICP4, and LAT (Fig. 1D). A closer analysis of ICP4 peaks demonstrated that the location and number of high confidence occupied sites did not alter significantly throughout infection (Fig. 1E). Instead the amount of ICP4 bound between distinct peaks decreased as genome number increased (Fig. 1C-D).

Although ICP4 exhibited a dense binding pattern at early times (2 hpi), with relatively broad, overlapping peaks we were able to determine high confidence binding sites. The final sites were consistent between biological replicates. We analyzed 100 bp extensions from the summits of the peaks seen at all 3 times (Fig. 1E) for motif discovery. DTSGKBDTBNHSG was the only motif discovered (Fig. 1B), where D is A, G, T; S is C or G; K is G or T; B is C, G, T; H is A, C, T. This binding motif was very close to that previously discovered using *in*

vitro techniques, RTCGTCNNYNYSG (DiDonato et al., 1991). These minor deviations may be due to protein partners or DNA binding proteins altering the binding ability of ICP4 *in vivo*.

These results demonstrate that ICP4 binds to specific sites, but also coats the genome early in infection forming a type of nucleoprotein. Due to mass action, ICP4-nucleoprotein changes as infection proceeds, limiting binding to predominantly strong cognate binding sites as the number of genomes increase due to replication.

ICP4 stabilizes GTF binding promoting cooperative preinitiation complex (PIC) assembly.

We wanted to investigate how the formation of ICP4-nucleoprotein affects the transcription factor landscape across the viral genome. We compared the binding of ICP4, Pol II, TATA-binding protein (TBP), SP1, Med1, and Med23 in ICP4 null (n12) and wild-type (WT) HSV-1 infected human fibroblasts at 2.5 hpi by ChIP-Seq (Fig. S1). In the absence of ICP4, we observed a decrease in binding for all the factors to most viral promoters, with the exception of IE genes (Fig. 2A-B), where there was an increase in binding. There were detectable, although highly reduced, peaks of TBP and SP1 binding to the UL23 promoter in the absence of ICP4. It has been previously shown that these sites are functional in the n12 background reflecting the basal binding activity of TBP and SP1 (Imbalzano et al., 1991) (Fig. 2B). Similar to UL23 we observed TBP and SP1 bound to select E promoters in the absence of ICP4, namely UL23, UL29, UL39, and UL50 (Fig. S1).

Med1 and Med23 bound the viral genome with an almost identical pattern (Fig. S1, Fig. 2), indicating they are parts of the Mediator complex bound early during viral infection. In WT infected cells, the binding of Mediator concentrated near the starts sites of ICP4-induced viral genes. However, the Mediator complex also densely coated the viral genome, resembling the ICP4 binding pattern. This dense coating was completely absent in n12

infection, demonstrating this phenotype is not an artifact of the IP. We suspect this reflects the fact that ICP4 and Mediator interact.

In the absence of ICP4, the binding of Pol II was reduced the most compared to the other transcription factors (Fig. 2A). This magnified difference is likely a result of the cooperative nature of Pol II recruitment requiring multiple protein-protein interactions. In summary ICP4 was required for robust recruitment of all GTF's tested, cooperatively recruiting Pol II to E and L promoters. The difference between n12 and WT shows the extent by which ICP4 mediated recruitment and bolstered the frequency of PIC assembly. IE promoters retained robust GTF recruitment via an independent mechanism involving a complex consisting of Oct-1, HCF and VP16 binding to TAATGARAT promoter elements (Preston et al., 1988; Stern and Herr, 1991; Stern et al., 1989)

Genome bound ICP4 does not affect accessibility.

Part of the mechanism of ICP4 action in the recruitment of GTFs to the genome may involve a role in the exclusion of repressive chromatin. To address this hypothesis, we investigated the relationship between presence of ICP4, the abundance of histones, and the accessibility of the genome. We used ChIP-Seq to compare the binding of ICP4, Pol II, and histone H3 in n12 and WT HSV-1 infected human fibroblasts at 2 hpi (Fig. 3). We found in both WT and n12 infection that the number of H3 reads mapped to the viral genome was 100-fold less than ICP4, and the pattern was nearly identical to input reads (Fig. 3A), with R^2 correlations of 0.0004 and 0.02 (Fig S2). These data demonstrated H3 binding to the viral genome was minimal and not reproducible. Furthermore, H3 binding was still minimal in the absence of ICP4 (n12). This was not due to technical issues as the number and quality of H3 reads mapped to the cellular genome for the same samples was approximately 10 million

with R^2 correlations of ≥0.97 (Fig. S3). We saw a similar trend with H3K4me3, H3K27ac, H3K9me3, and H3K27me3 reads mapped to the viral genome (Fig S2&4).

Although H3 binding to the viral genome was similar in WT and n12 infection, we could not rule out the role of an alternative protein occluding the genome. To investigate genomic accessibility, we performed ATAC-Seq. Human fibroblasts were infected with WT and n12 HSV-1 at an MOI of 10 pfu/cell and collected prior to the onset of genome replication. Quantification of ChIP-Seq input reads allowed us to determine that the approximate number of genomes per cell in WT and n12 infection was 169 and 254, respectively (Fig 3C). This value is consistent with infecting at an MOI of 10 pfu/cell and an approximately particle to pfu ratio of 20-30. We normalized ATAC-Seq traces to adjust for sequencing depth and input genome number. We observed even tagmentation in both conditions absent the nucleosomal laddering visible on the cellular genome (Fig 3B). Quantification of ATAC-Seq reads determined that the viral genome in n12 and WT was 2.8 and 4-fold more accessible than the cellular genome (Fig 3C). As we harvested samples pre-replication, we expect that a significant portion of viral genomes are defective and will not undergo replication. Our ATAC-Seq data is thus an average of tagmentation for defective and active viral genomes. For this reason, we expect our accessibility calculation is an underestimate. We conclude that the viral genome was much more accessible than the cellular genome, and this increased accessibility was not ICP4-dependent. ICP4 binding and GTF recruitment, not viral genome accessibility, was responsible for robust GTF binding.

ICP4 binds to cellular transcription start sites (TSS) early during infection.

Immunofluorescence (IF) studies of HSV-1 infection depict colocalization of ICP4 with EdC-labeled viral genomes and exclusion from dense areas of cellular chromatin

193 (Dembowski and DeLuca, 2015). This phenomenon is so well established that ICP4 is largely
194 used in IF studies as a proxy for HSV-1 genomes. To ascertain if ICP4 also binds to the
195 cellular genome, we aligned our ICP4 ChIP-Seq data from 2, 4, and 6 hpi to the cellular
196 genome. ICP4 bound to the cellular genome in a manner quite distinctive from the pattern
197 observed on the viral genome. ICP4 only bound in distinct peaks around cellular transcription
198 start sites (TSS) (Fig. 4E-F) of a subset of cellular genes (Fig. 4B). These genes grouped
199 ontologically to common housekeeping functions including pathways related to chromatin,
200 transcription, and metabolism (Fig 4C). This binding reduced from 2 to 4 hpi, and become
201 negligible at 6 hpi (Fig. 4). At 2 hpi ICP4 bound to the cellular genome at 5,727 sites (Fig. 4D)
202 or 0.002 peaks per kbp, whereas ICP4 bound to the viral genome at 122 sites (Fig. 1E) or 0.8
203 peaks per kbp. Similar we found a much greater density of ICP4 binding motifs present in the
204 viral genome (2 motifs/kbp) than the cellular genome (0.02 motifs/kbp). We observed ICP4
205 binding peaks that did not localize at an ICP4 binding consensus (Fig. 4A) suggesting that
206 ICP4 may associate with the cellular genome by an alternative mechanism. We conclude that
207 ICP4 bound to the cellular genome early during infection, when the relative concentration of
208 ICP4 to viral genomes is still quite high. The amount of ICP4 on the cellular genome quickly
209 dropped off as viral genome number increased and ICP4 preferentially bound to the viral
210 genome.
211

212 **ICP4 binding is restricted to accessible regions of the cellular genome.**

213 Since ICP4 bound to a subset of cellular genes near mRNA start sites (Fig. 4), We
214 hypothesized that ICP4 only bound to accessible regions of the cellular genome. To test this
215 hypothesis, we performed ChIP-Seq for ICP4, Pol II, Histone H3 (H3), euchromatic markers
216 H3K4-trimethyl (H3K4me3) and H3K27-acetyl (H3K27ac), and heterochromatic marker

H3K9-trimethyl (H3K9me3) and H3K27-trimethyl (H3K27me3) on MRC5 cells that were infected with HSV for 2 h. Cellular TSS were stratified using k-means clustering as high and low ICP4 binding (Fig. 5A). TSS with high ICP4 binding were also bound by Pol II and adjacent to euchromatic markers. TSS with low ICP4 binding were associated with only heterochromatic markers. Furthermore, genes clustered as high ICP4 binding had higher tagmentation frequency when assessed using ATAC-Seq (Fig. 5 B). The data was mapped for representative cellular genes in Fig. S5. We quantified the relationship between ICP4 and cellular chromatin in Fig. 5C-D. We found that the binding pattern of ICP4 was directly related (Spearman coefficient ≥0.5) to Pol II and cellular euchromatin, clustering as most similar to H3K27ac and H3K4me3 (Fig. 5C). The heterochromatic markers, H3K9me3 and H3K27me3, clustered together, and were not correlated (Spearman coefficient ~0) to ICP4 or cellular chromatin. These results were corroborated by analysis of distinct peaks called using MACS (Fig. 5D). Interestingly, ICP4 bound regions had little overlap with their cognate binding motifs (Fig. 5D). A closer analysis of the actual genomic region where each factor bound, revealed that 82% of ICP4 bound regions were within 1 kb of a promoter (Fig. S6). By comparison only 10% of ICP4 predicted binding motifs were within 1 kb of a promoter. Furthermore, the euchromatic regions of the cellular genome that were occupied by ICP4 in infected cells were also euchromatic in uninfected cells, indicating that ICP4 does not globally promote open chromatin in these regions of the genome (Fig. 5A-B). These data support a model in which ICP4 is able to bind nonspecifically to accessible regions of the cellular genome, namely active promoters, early in infection when the relative concentration of ICP4 is high.

ICP4 mediates depletion of Pol II on cellular promoters

We observed depletion in Pol II binding to cellular promoters with infection (Fig. 5A). This observation is consistent with prior studies, which assessed HSV-1 infection post-

241 replication at 3, 4, or 6 hpi (Abrisch et al., 2016; Birkenheuer et al., 2018; McSwiggen et al.,
242 2019). As we harvested samples prior to the onset of genome replication (2 hpi) we
243 hypothesized that ICP4, which is produced immediately upon viral infection was responsible.
244 First, we determined the effect of ICP4 on cellular promoters before the onset of genome
245 replication. We mock-infected or infected fibroblasts at 10 pfu/cell with WT or ICP4-null (n12)
246 HSV-1 for 2 h. We chose this early time point to ensure the effect we see is due to the
247 absence of ICP4, rather than an E or L viral gene product which cannot be produced in the
248 absence of ICP4. We observed depletion of Pol II occupancy on cellular mRNA promoters
249 only in WT infection (Fig. 6A-B). Thus we concluded that ICP4 was required for depletion of
250 Pol II from host mRNA promoters, and this effect was independent of viral genome copy
251 number.

252 We then assessed whether ICP4 was continuously required for cellular Pol II
253 depletion, namely if ICP4 was still essential even after the onset of genome replication. We
254 used a temperature sensitive ICP4 mutant (tsKos), in which growth at nonpermissive
255 temperature (39.6^{o}C) results in loss of ICP4 in the nucleus (Dremel & DeLuca 2019). We
256 infected fibroblasts with tsKos grown at permissive conditions (P), shifted up from permissive
257 to nonpermissive conditions at 4 hpi (S), or nonpermissive conditions (N). In this system we
258 can separate the role of ICP4 in Pol II depletion, from ICP4's requirement in E and L
259 transcription and viral genome replication. Infected cells were harvested at 4 or 6 hpi and Pol
260 II ChIP-Seq was performed. We used nonpermissive conditions as a surrogate to mock-
261 infection, as we just established that Pol II depletion does not occur in n12 infection (Fig. 6A-
262 B). We observed significant depletion of Pol II from cellular promoters in permissive and
263 shifted samples (Fig. 6C-D). Pol II depletion was not directly related to viral genome copy
264 number. tsKos shifted up had the highest number of viral genomes present, but did not reach

265 the same level of cellular Pol II depletion as cells grown at permissive temperature for the
266 same length of time. These data suggest that the viral genome is not solely responsible for
267 preferential recruitment of cellular Pol II. Instead ICP4 bound to the viral genome is required
268 for depletion of Pol II from cellular promoters. These results suggest a model in which
269 genome replication facilitates host Pol II depletion when the relative number of ICP4 to viral
270 genomes is high (2h). As the number of ICP4 bound viral genomes increased, we observed a
271 corresponding decrease in Pol II on host promoters.

DISCUSSION

ICP4 as a sink for general transcription factors

ICP4 is synthesized shortly after the viral genome enters the nucleus and remains associated with the genome through all phases of infection. Our data demonstrated that ICP4 bound promiscuously to the viral genome prior to DNA replication. At this time point, ICP4 was present at a relatively high concentration which likely promoted multimerization on DNA through ICP4-ICP4 interactions (Kuddus and DeLuca, 2007). We observed a similar phenotype for ICP4's interaction partner, Mediator (Lester and Deluca; Wagner and DeLuca, 2013). Components of Mediator bound generally to the viral genome, concentrating near viral TATA boxes. Additional protein-Mediator interactions likely contribute to this distribution. This is a unique recruitment phenotype for Mediator which binds exclusively at cellular TSS via multiple protein-protein interactions. In the absence of ICP4, these interactions were not sufficient to support Mediator binding to the viral genome. With the exception of Mediator recruitment to IE promoters which does not require ICP4 and reflects the activity of VP16 (Batterson and Roizman, 1983; Campbell et al., 1984). Similarly, we observed a 2 to 10-fold decrease in recruitment of Pol II, TBP, and Sp1 to viral E and L promoters without ICP4. This minimal level of recruitment is insufficient to support transcription, which explains why only IE transcripts are efficiently transcribed in the absence of ICP4.

ICP4-dependent GTF recruitment was not due to a global accessibility change. In the absence of ICP4 the viral genome remained absent of histones and had little change in tagmentation frequency. This is most likely due to the action of ICP0, which is an IE protein expressed in the absence of ICP4 and has been shown to preclude histones from the genome (Cliffe and Knipe, 2008; Ferenczy and DeLuca, 2009). We posit that ICP4's ability to interact and recruit Mediator and TFIID generally to the viral genome creates a local concentration gradient. Ultimately this increases the incidence of Pol II transcription

297 machinery recruitment to the viral genome, which is stabilized by contact with additional
298 protein-DNA, protein-protein interactions. These data demonstrated the critical role ICP4
299 serves as a general viral transcription factor, essential for activation and continued
300 transcription of E and L genes.
301
302 **ICP4 differentiates between the viral and cellular genome**
303 ICP4 possesses the ability to bind to double stranded DNA independent of sequence,
304 an ability that is facilitated by ICP4 oligomerization on the genome. At early time points, when
305 the relative concentration of ICP4 to the viral genome was high, we observed promiscuous
306 ICP4 binding. This coating phenotype provides an explanation for why no specific binding
307 sites on the genome affect the ability of ICP4 to activate transcription (Coen et al., 1986;
308 Smiley et al., 1992). Instead the high density of ICP4 binding motifs on the viral genome
309 aggregate to create a global affinity for ICP4.
310 Early during infection, we also observed binding of ICP4 to cellular promoters, a novel
311 observation. ICP4 only bound highly transcribed cellular promoters—largely housekeeping
312 genes—and specifically bound where there was an absence of histones, adjacent to
313 euchromatic markers. The consequences, if any, of this binding for the transcription of
314 specific cellular genes remains to be determine. The binding of ICP4 to the cellular genome
315 was greatly diminished by 4 hpi, which corresponded to 3-4 viral genome duplications. At this
316 time point we also observed a decrease in ICP4 coating the viral genome. However, ICP4 still
317 bound abundantly, concentrating adjacent to strong cognate binding sites. This is most likely
318 due to replication of the viral genome producing more ICP4 binding targets. Simple mass
319 action results in binding to predominantly higher affinity sites. We propose that the binding
320 preference of ICP4 for the viral genome is due to the 100-fold higher density of cognate
321 binding sites and absence of cellular histones.

ICP4 as both viral transcription factor and chromatin

HSV-1 productive infection generates 1,000-10,000 viral progeny per infected cell within a 24 hour window. To facilitate this rampant transcriptional shift HSV-1 manipulates host Pol II machinery to prioritize viral mRNAs. By 6 hpi viral mRNA's comprise almost 50% of the total mRNA present in the host nucleus (Dremel and DeLuca, 2019). Furthermore, binding of Pol II to cellular promoters dramatically decreases upon HSV-1 infection (Abrisch et al., 2016; Birkenheuer et al., 2018). A recent study concluded that viral replication compartments efficiently enrich Pol II into membraneless domains (McSwiggen et al., 2019). Herein we identified the viral factor responsible for coopting the host Pol II machinery.

McSwiggen et al. proposed this phenomenon was dependent on the absence of nucleosomes which made the viral genome 100-fold more accessible than the cellular genome. While we agree that this accessibility is critical for viral infection, we believe it is essential for ICP4 binding. Similar to cellular chromatin, ICP4 coats the viral genome throughout productive infection (Fig. 7). However, ICP4 also functions to scaffold Pol II transcription machinery to the viral genome. We demonstrated that Pol II depletion from cellular promoters was dependent on the number of ICP4 bound viral genomes. We propose that one or more components of the PIC, such as the ICP4 binding partners TFIID and Mediator, are limiting and ICP4 recruits these factors to the viral genome. As the number of viral genomes bound by ICP4 increases, the limiting PIC components no longer contacts cellular promoters. Ultimately this results in decreased Pol II occupancy on host promoters, preventing cellular transcription. This mechanism is essential to facilitate the rapidly progressing infection while limiting the extent to which the host can respond to viral challenge. This mechanism may explain how HSV-1 and -2 complete the infectious life cycle faster than other herpesviruses.

MATERIALS AND METHODS

Cells and Viruses

Vero (African green monkey kidney) and MRC5 (human fetal lung) cells were obtained from and propagated as recommended by ATCC. Viruses used in this study include n12 (DeLuca and Schaffer, 1988), tsKos (Dremel and DeLuca, 2019) and KOS. n12 virus stocks were prepared and titered in a Vero-based ICP4 complementing cell line, E5. KOS virus stocks were prepared and titered in Vero cells. tsKos virus stocks were prepared and tittered in Vero cells at permissive temperature (33.5^{o}C).

Antibodies

The following antibodies were used: Pol II 4H8 (AbCam #ab5408), TBP (AbCam #ab51841), Sp1 (SantaCruz #sc-17824), Med1 (BD Pharmingen #550429), Med23 (Bethyl #A300-793A), H3K4me3 (Abcam #ab12209), H3K27me3 (AbCam #ab6002), H3K27acetyl (AbCam #ab4729), H3K9me3 (AbCam #ab176916), H3 (AbCam #1791), and ICP4 58S (derived from hybridomas-ATCC HB8183).

Viral Infection

MRC5 cells were infected with 10 PFU per cell. Virus was adsorbed in tricine-buffered saline (TBS) for 1 hour at room temperature. Viral inoculum was removed, and cells were washed quickly with TBS before adding 2% FBS media. Infected samples were incubated at 37^{o}C unless otherwise specified.

ChIP-Sequencing

Infected cells were treated with 5 mL of 25% formaldehyde for 15 minutes at room temperature, followed by 5 mL of 2.5 M glycine. All following steps were performed at 4^{o}C unless otherwise stated. Cultures were washed with TBS and scraped into 50 mL of FLB [5 mM 1,4-Piperazinediethanesulfonic acid (PIPES) pH 8, 85 mM KCl, 0.5% Igepal CA-630, 1x Roche protease inhibitor cocktail]. Cells were pelleted by low-speed centrifugation,

resuspended in 1.1 mL RIPA buffer [1x phosphate-buffered saline (PBS), 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), 1x Roche protease inhibitor cocktail]. Sample was sonicated for 6 intervals of 30 seconds with a Sonics Vibra-Cell VCX 130 sonicator equipped with a 3-mm microprobe and pelleted at 2000 xg for 15 minutes. 50 µl was stored as an input control, and the remainder was divided equally to use in immunoprecipitations (IP). 2-4x10^7 MRC5 cells were applied per IP. Samples were immunoprecipitated with 25 µg (TBP, Sp1), or 10 µg (Pol II, H3K4me3, H3K27me3, H3K27ac, H3K9me3, H3) antibody. Antibody was previously bound to 50 µL of Dynabeads M280 sheep anti-mouse IgG beads, or Dynabeads M280 sheep anti-rabbit IgG beads in 5% bovine serum albumin (BSA) 1x PBS overnight. DNA samples were bound to the antibody-bead complex overnight rotating. The IP mixtures were washed seven times with LiCl wash buffer [100 mM Tris-HCl buffer pH 7.5, 500 mM LiCl, 1% Igepal CA-630, 1% sodium deoxycholate] and once with Tris-EDTA (TE) buffer. Beads were resuspended in IP elution buffer [1% SDS, 0.1M $NaHCO_3$] and incubated at 65^{o}C for 2 h 900 rpm. Input aliquot was suspended in IP elution buffer. Input and IP samples were incubated at 65^{o}C 900 rpm overnight. The samples were extracted with phenol-chloroform-isoamyl alcohol (25:24:1) and with chloroform-isoamyl alcohol (24:1) and then purified using Qiagen PCR cleanup columns. Each sample was quantified using a Qubit 2.0 fluorometer (Invitrogen) and 2-20 ng was used to create sequencing libraries using the NEBNext Ultra II DNA Library preparation kit (NEB #E7103S). Libraries were quantified using the Agilent DNA 7500 Kit, and samples were mixed together at equimolar concentration. Illumina HiSeq 2500 single-end 50 bp sequencing was carried out at the Tufts University Core Facility.

ATAC-Sequencing

We adapted the protocol from Buenrostro et al.(Buenrostro et al., 2013). Briefly, 2 million MRC5 cells were plated into 60 mm dishes and allowed to grow overnight. Cells were

infected as described above. Uninfected and n12 infected cells were harvested at 4 hpi. WT HSV-1 infected cells were harvested pre-replication at 2 hpi. Infected samples were washed once with chilled TBS and lysis-1 buffer [10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$]. Samples were incubated with 2 mL lysis-2 buffer [10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Igepal CA-630] for 3 minutes on ice. Cells were gently resuspended and dounced until nuclei were visible via trypan blue staining. Nuclei were spun at 500 g for 10 min at 4^{o}C and resuspended in lysis-1 buffer. 500 µL (10^6 cells) was transferred to an epindorf tube and spun at 500 g for 10 min at 4^{o}C. Nuclei were resuspended in 22 µL buffer TD (Illumina Catalog No. 15027866) 2.5µL TDE1 (Illumina Catalog No. 15027865) and 22.5 µL water and incubated at 37^{o}C for 30 min gently shaking. DNA was purified using the MinElute PCR purification kit (Qiagen Cat No./ID: 28004). PCR amplification was performed for 8-14 total cycles. Libraries were quantified using the Agilent DNA 7500 Kit, and samples were mixed together at equimolar concentration. Illumina HiSeq 2500 single-end 50 bp sequencing was carried out at the Tufts University Core Facility.

Data Analysis

ChIP-Seq

Data was uploaded to the Galaxy web platform, and we used the public server at usegalaxy.org to analyze the data (Guerler et al., 2018). Data was first aligned using Bowtie2 (Langmead and Salzberg, 2012) to the human genome (hg38), and then unaligned reads were mapped to the HSV-1 strain KOS genome (KT899744.1). Bam files were visualized using DeepTools bamcoverage (Ramírez et al., 2016) with a bin size of 1 to generate bigwig files. Data was viewed in IGV viewer and exported as EPS files. Bigwig files were normalized for sequencing depth and genome quantity. Mapped reads were multiplied by the “norm factor” which was calculated as the inverse of $(Input\ cellular\ reads)/(Input\ cellular\ +$

$viral\ mapped\ reads\ (TMR))\times Billion\ sample\ TMR$ or $(Input\ viral\ reads\ )/TMR\times Million\ sample\ TMR$. ChIP-Seq experiments were repeated for a total of 2 to 4 biological replicates. The normalized bigwig files were averaged between replicates. Heatmaps and gene profiles were generated using MultiBigwigSummary (Ramírez et al., 2016) on normalized cellular bigwig files to all UCSC annotated mRNAs. Gene profiles and heatmaps were plotted using plotProfile and plotHeatmap (Ramírez et al., 2016). Spearman correlation analysis was performed using deeptools plotCorrelation on multiBigwigSummary limited to cellular transcripts (Ramírez et al., 2016).

Peak Calling

Viral peaks were called using MACS2 call peak (Feng et al., 2012), pooling treatment and control files for each condition. Due to the small size of the viral genome (151974 bp) we could not use the shifting model option (--nomodel). To offset the dense binding of ICP4 we used a fixed background lambda as local lambda for every peak region and a more sophisticated signal processing approach to find subpeak summits in each enriched peak region (--call-summits).

Cellular peaks were called using MACS2 (Feng et al., 2012). We first removed non-uniquely mapped sequences with SAMtools, filter SAM or BAM for a minimum MAPQ quality score of 20 (Subgroup et al., 2009). We determined the approximate extension size for each IP using MACS2 predictd, and averaging the size estimate between replicates. We ran MACS2 call peak for individual replicates and pooled samples with no shifting model (--nomodel). To determine high confidence peaks present in each MACS2 output we used Galaxy Operate on Genomic Intervals, Join. Peak intersection was analyzed for intersection size and jaccard statistic using JaccardBed (Ramírez et al., 2016). ChIPseeker was run on MACS2 outputs to assess the cellular regions bound in each condition (Yu et al., 2015).

Motif Discovery

Bedtools Multiple Intersect (Quinlan and Hall, 2010) was used to compare the MACS2 output for ICP4 IP at 2, 4, and 6 h. A BED file was generated for regions +/- 100 bp from the summits of each identified peak. Peaks in common between all three experimental conditions were used to generate a fasta file using GetFastaBed (Quinlan and Hall, 2010) Peaks present in all three time points were submitted to MEME v.4.11.1.0 for motif analysis (Bailey et al., 2009). The consensus sequence in Fig. 1 had the most significant E-value, and was the only motif found in more than 5 peaks.

Correlation Analysis

To assess quality and reproducibility of data we assessed normalized bigwig files for each IP replicate. For cellular and viral alignments we ran MultiBigwigSummary (Ramírez et al., 2016) with a bin size of 10,000 and 50 bp, respectively. Raw bin counts were plotted and a linear regression analysis was performed (Fig. S2-3).

ATAC-Seq

Data was first aligned using Bowtie2 (Langmead and Salzberg, 2012) to the human genome (hg38), and then unaligned reads were mapped to the HSV-1 strain KOS genome (KT899744.1). Bam files were visualized using DeepTools bamcoverage (Ramírez et al., 2016) with a bin size of 1 to generate bigwig files. Data was viewed in IGV viewer and exported as EPS files. Bigwig files were normalized for sequencing depth, or billion total reads. Heatmaps and gene profiles were generated using MultiBigwigSummary (Ramírez et al., 2016) on normalized cellular bigwig files to all UCSC annotated mRNAs. Gene profiles and heatmaps were plotted using plotProfile and plotHeatmap (Ramírez et al., 2016). To calculate the percentage of total DNA corresponding to the virus or host in n12 and WT HSV-1 infection, we utilized ChIP-Seq input reads. We calculated the average percentage of total reads which mapped to either the virus or host in four biological replication ChIP-Seq

470 samples. We used this value to calculate the number of viral genomes contained within each
471 nucleus. This value was used to determine the tagmentation enrichment observed relative to
472 the actual amount of genome content present.

473 **Data Availability**

474 All data are publicly accessible in the SRA database (PRJNA553543, PRJNA553555,
475 PRJNA553559, PRJNA553563, PRJNA508791).

476

REFERENCES

Abrisch, R.G., Eidem, T.M., Yakovchuk, P., Kugel, J.F., and Goodrich, J.A. (2016). Correction for Abrisch et al., Infection by Herpes Simplex Virus 1 Causes Near-Complete Loss of RNA Polymerase II Occupancy on the Host Cell Genome. J Virol *90*, 11279.

Alwine, J.C., Steinhart, W.L., and Hill, C.W. (1974). Transcription of herpes simplex type 1 DNA in nuclei isolated from infected HEp-2 and KB cells. Virology *60*, 302-307.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic acids research *37*, W202-W208.

Batterson, W., and Roizman, B. (1983). Characterization of the herpes simplex virion-associated factor responsible for the induction of alpha genes. J Virol *46*, 371-377.

Birkenheuer, C.H., Danko, C.G., and Baines, J.D. (2018). Herpes Simplex Virus 1 Dramatically Alters Loading and Positioning of RNA Polymerase II on Host Genes Early in Infection. J Virol *92*.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods *10*, 1213.

Campbell, M.E., Palfreyman, J.W., and Preston, C.M. (1984). Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. Journal of Molecular Biology *180*, 1-19.

Carrozza, M.J., and DeLuca, N.A. (1996). Interaction of the viral activator protein ICP4 with TFIID through TAF250. Mol Cell Biol *16*, 3085-3093.

Cliffe, A.R., and Knipe, D.M. (2008). Herpes simplex virus ICP0 promotes both histone removal and acetylation on viral DNA during lytic infection. J Virol *82*, 12030-12038.

Coen, D.M., Weinheimer, S.P., and McKnight, S.L. (1986). A genetic approach to promoter recognition during trans induction of viral gene expression. Science *234*, 53-59.

Courtney, R.J., and Benyesh-Melnick, M. (1974). Isolation and characterization of a large molecular-weight polypeptide of herpes simplex virus type 1. Virology *62*, 539-551.

DeLuca, N.A., McCarthy, A.M., and Schaffer, P.A. (1985). Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. Journal of Virology *56*, 558-570.

DeLuca, N.A., and Schaffer, P.A. (1988). Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. J Virol *62*, 732-743.

Dembowski, J.A., and DeLuca, N.A. (2015). Selective recruitment of nuclear factors to productively replicating herpes simplex virus genomes. PLoS Pathog *11*, e1004939.

Dembowski, J.A., and DeLuca, N.A. (2018a). Temporal viral genome-protein interactions define distinct stages of productive herpesviral infection. MBio *9*, e01182-01118.

Dembowski, J.A., and DeLuca, N.A. (2018b). Temporal Viral Genome-Protein Interactions Define Distinct Stages of Productive Herpesviral Infection. MBio *9*.

Dembowski, J.A., Dremel, S.E., and DeLuca, N.A. (2017). Replication-coupled recruitment of viral and cellular factors to herpes simplex virus type 1 replication forks for the maintenance and expression of viral genomes. PLoS Pathog *13*, e1006166.

DiDonato, J.A., Spitzner, J.R., and Muller, M.T. (1991). A predictive model for DNA recognition by the herpes simplex virus protein ICP4. J Mol Biol *219*, 451-470.

Dixon, R.A., and Schaffer, P.A. (1980). Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. J Virol *36*, 189.

525 Dremel, S.E., and DeLuca, N.A. (2019). Genome replication affects transcription factor
526 binding mediating the cascade of herpes simplex virus transcription. Proc Natl Acad Sci
527 U S A *116*, 3734-3739.
528 Everett, R.D., Sourvinos, G., and Orr, A. (2003). Recruitment of herpes simplex virus type 1
529 transcriptional regulatory protein ICP4 into foci juxtaposed to ND10 in live, infected
530 cells. J Virol *77*, 3680-3689.
531 Faber, S.W., and Wilcox, K.W. (1986). Association of the herpes simplex virus regulatory
532 protein ICP4 with specific nucleotide sequences in DNA. Nucleic Acids Res *14*, 6067-
533 6083.
534 Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment
535 using MACS. Nature protocols *7*, 1728.
536 Ferenczy, M.W., and DeLuca, N.A. (2009). Epigenetic modulation of gene expression from
537 quiescent herpes simplex virus genomes. J Virol *83*, 8514-8524.
538 Guerler, A., Baker, D., Clements, D., Afgan, E., Taylor, J., Batut, B., Rasche, H., Grüning,
539 B.A., van den Beek, M., Nekrutenko, A.*, et al.* (2018). The Galaxy platform for
540 accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic
541 acids research *46*, W537-W544.
542 Honess, R.W., and Roizman, B. (1974a). Regulation of herpesvirus macromolecular
543 synthesis. J Virol *14*, 8-19.
544 Honess, R.W., and Roizman, B. (1974b). Regulation of herpesvirus macromolecular
545 synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J
546 Virol *14*, 8-19.
547 Honess, R.W., and Roizman, B. (1975). Regulation of herpesvirus macromolecular synthesis:
548 sequential transition of polypeptide synthesis requires functional viral polypeptides.
549 Proc Natl Acad Sci USA *72*, 1276-1280.

Imbalzano, A.N., Coen, D.M., and DeLuca, N.A. (1991). Herpes simplex virus transactivator ICP4 operationally substitutes for the cellular transcription factor Sp1 for efficient expression of the viral thymidine kinase gene. J Virol *65*, 565-574.

Knipe, D.M., and Cliffe, A. (2008). Chromatin control of herpes simplex virus lytic and latent infection. Nat Rev Microbiol *6*, 211-221.

Kuddus, R.H., and DeLuca, N.A. (2007). DNA-dependent oligomerization of herpes simplex virus type 1 regulatory protein ICP4. J Virol *81*, 9230-9237.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357.

Lester, J.T., and Deluca, N.A. Herpes Simplex Virus 1 ICP4 Forms Complexes with TFIID and Mediator in Virus-Infected Cells. J Virol *85*, 5733-5744.

Liang, Y., Vogel, J.L., Narayanan, A., Peng, H., and Kristie, T.M. (2009). Inhibition of the histone demethylase LSD1 blocks alpha-herpesvirus lytic replication and reactivation from latency. Nat Med *15*, 1312-1317.

McGeoch, D.J., Dalrymple, M.A., Davison, A.J., Dolan, A., Frame, M.C., McNab, D., Perry, L.J., Scott, J.E., and Taylor, P. (1988). The complete DNA sequence of the long unique region in the genome of

herpes simplex virus type 1. J Gen Virol *69*, 1531-1574.

McGeoch, D.J., Dolan, A., Donald, S., and Brauer, D.H. (1986). Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res *14*, 1727-1745.

McGeoch, D.J., Dolan, A., Donald, S., and Rixon, F.J. (1985). Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J Mol Biol *181*, 1-13.

McSwiggen, D.T., Hansen, A.S., Teves, S.S., Marie-Nelly, H., Hao, Y., Heckert, A.B., Umemoto, K.K., Dugast-Darzacq, C., Tjian, R., and Darzacq, X. (2019). Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. Elife *8*.

Powell, K.L., and Purifoy, D.J. (1976). DNA-binding proteins of cells infected by herpes simplex virus type 1 and type 2. Intervirology *7*, 225-239.

Preston, C.M. (1979). Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. J Virol *29*, 275-284.

Preston, C.M., Frame, M.C., and Campbell, M.E. (1988). A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell *52*, 425-434.

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

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-W165.

Sampath, P., and Deluca, N.A. (2008). Binding of ICP4, TATA-binding protein, and RNA polymerase II to herpes simplex virus type 1 immediate-early, early, and late promoters in virus-infected cells. J Virol *82*, 2339-2349.

Smiley, J.R., Johnson, D.C., Pizer, L.I., and Everett, R.D. (1992). The ICP4 binding sites in the herpes simplex virus type 1 glycoprotein D (gD) promoter are not essential for efficient gD transcription during virus infection. J Virol *66*, 623-631.

Stern, S., and Herr, W. (1991). The herpes simplex virus trans-activator VP16 recognizes the Oct-1 homeo domain: evidence for a homeo domain recognition subdomain. Genes Dev *5*, 2555-2566.

598 Stern, S., Tanaka, M., and Herr, W. (1989). The Oct-1 homoeodomain directs formation of a
599 multiprotein-DNA complex with the HSV transactivator VP16. Nature *341*, 624-630.
600 Subgroup, G.P.D.P., Wysoker, A., Handsaker, B., Marth, G., Abecasis, G., Li, H., Ruan, J.,
601 Homer, N., Durbin, R., and Fennell, T. (2009). The Sequence Alignment/Map format
602 and SAMtools. Bioinformatics *25*, 2078-2079.
603 Wagner, L.M., and DeLuca, N.A. (2013). Temporal association of herpes simplex virus ICP4
604 with cellular complexes functioning at multiple steps in Pol II transcription. PLOS ONE
605 *8*, e78242.
606 Watson, R.J., and Clements, J.B. (1980). A herpes simplex virus type 1 function continuously
607 required for early and late virus RNA synthesis. Nature *285*, 329.
608 Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP
609 peak annotation, comparison and visualization. Bioinformatics *31*, 2382-2383.
610

FIGURES LEGENDS

Fig. 1. ICP4 binding at key points in the viral life cycle. MRC5 cells were infected with HSV-1 for 2, 4, or 6 h, and ChIP-Seq for ICP4 was performed. A) Quantification of ICP4 binding (IP), viral genomes (Input), or ICP4 binding per genome (IP/Input). B) ICP4 consensus binding motif. C-D) ICP4 binding normalized per viral genome. Viral ORFs are indicated, color coded by gene class with IE as yellow, E as green, leaky late (L1) as blue, and true late (L2) as purple. FIMO identified genome sequences matching the consensus motif in B are indicated in red. E) Intersection of MACS2 identified ICP4 occupied regions.

Fig. 2. ICP4 recruitment of host Pol II machinery to viral promoters. MRC5 cells were infected with an ICP4 null mutant (n12) or HSV-1 (WT) for 2.5 h and ChIP-Seq for ICP4, Pol II, TBP, Sp1, Med1, and Med23 was performed. Data quantified as in Fig. 1. A) Fold change of n12 over WT aligned to the viral genome. Loci with greater binding in n12 or WT are colored in green or red, respectively. B) ChIP-Seq reads normalized per viral genome and aligned to canonical IE and E genes.

Fig. 3. ICP4 dependence of viral genome accessibility and Histone H3 binding. MRC5 cells were infected with an ICP4 null mutant (n12) or HSV-1 (WT) and harvested prior to genome replication. A) ChIP-Seq for ICP4, Pol II, and H3 was performed. Data aligned to the viral genome and quantified as in Fig. 1. B-C) ATAC-Seq data normalized for total sequencing depth. C) Quantitative analysis of ATAC-Seq data, measuring the relative tagmentation enrichment for the virus or host as compared to expected.

Fig. 4. ICP4 binds cellular promoters during early infection. MRC5 cells were infected with HSV-1 for 2, 4, or 6 h, and ChIP-Seq for ICP4 was performed. Data was aligned to the

636 human genome (hg38) and normalized for cellular genome sampling. A) FIMO identified
637 genome sequences matching the consensus motif in Fig. 1B are indicated in red. B, E-F)
638 ICP4 binding to cellular mRNAs. Scaled from E) transcription start site (TSS) to transcription
639 end site (TES), or B, F) +/-1 kilobase from the TSS. D) Intersection of MACS2 peaks. C) Top
640 10 enriched Reactome groups for genes (n=2190) with ICP4 bound at the promoter at 2 h.
641
642 **Fig. 5. Association between cellular ICP4 binding and chromatin.** MRC5 cells were
643 uninfected or infected with HSV-1 for 2 h. Data was aligned to the human genome (hg38) and
644 normalized for sequencing depth. A, C-D) ChIP-Seq data for ICP4, Pol II, H3, H3K4me3,
645 H3K27acetyl, H3K9me3, and H3K27me3. B) ATAC-Seq data normalized as in Fig. 3. A-B)
646 Sequencing data centered +/-1 kilobase from the TSS of cellular mRNAs. Data was stratified
647 for ICP4 binding using K-means clustering. C) Spearman correlation analysis, limited to
648 cellular transcripts. D) Intersection of MACS2 peaks, analyzed as number of intersecting
649 peaks or Jaccard statistic.
650
651 **Fig. 6. The role of ICP4 in Pol II loss on host promoters.** MRC5 cells were A-B) uninfected
652 or infected with n12 or WT HSV-1 and harvested prior to genome replication or C-D) infected
653 with tsKos and grown at permissive conditions (P), shifted up from permissive to
654 nonopermissive conditions at 4 hpi (S), or nonpermissive conditions (N). A-D) ChIP-Seq for
655 Pol II was performed and data was aligned to cellular promoters +/- 1 kilobase from TSS. The
656 average signal for each condition plotted as a line graph.
657
658 **Fig. 7. Model for ICP4 function.** ICP4 preferentially binds to the more accessible viral
659 genome recruiting cellular transcription factors preferentially to the viral genome, thus
660 activating the virus and inhibiting cellular transcription.

Fig. S1. MRC5 cells were infected with an ICP4 null mutant (n12) or HSV-1 (WT) for 2.5 h and ChIP-Seq for ICP4, Pol II, TBP, Sp1, Med1, and Med23 was performed. ICP4 binding normalized per genome quantity. Viral ORFs are indicated, color coded by gene class with IE as yellow, E as green, L1 as blue, and L2 as purple.

Fig. S2. Analysis of ChIP-Seq data quality. Normalized viral aligned bigwig files were assessed using MultiBigwigSummary in 50 bp bins. The values within these bins were plotted for biological replicates and a linear regression analysis was performed.

Fig. S3. Analysis of ChIP-Seq data quality. Normalized cellular aligned bigwig files were assessed using MultiBigwigSummary in 10,000 bp bins. The values within these bins were plotted for biological replicates and a linear regression analysis was performed.

Fig. S4. MRC5 cells were infected with an ICP4 null mutant (n12) or HSV-1 (WT) for 2 h and ChIP-Seq for ICP4, Pol II, H3, H3K4me3, H3K27acetyl, H3K9me3, and H3K27me3 was performed. A-B) Factor binding normalized per genome quantity. Viral ORFs are indicated, color coded by gene class with IE as yellow, E as green, L1 as blue, and L2 as purple.

Fig. S5. MRC5 cells were uninfected or infected with HSV-1 for 2 h, and ChIP-Seq for ICP4, Pol II, H3, H3K4me3, H3K27acetyl, H3K9me3, and H3K27me3 was performed. Data was aligned to the human genome (hg38).

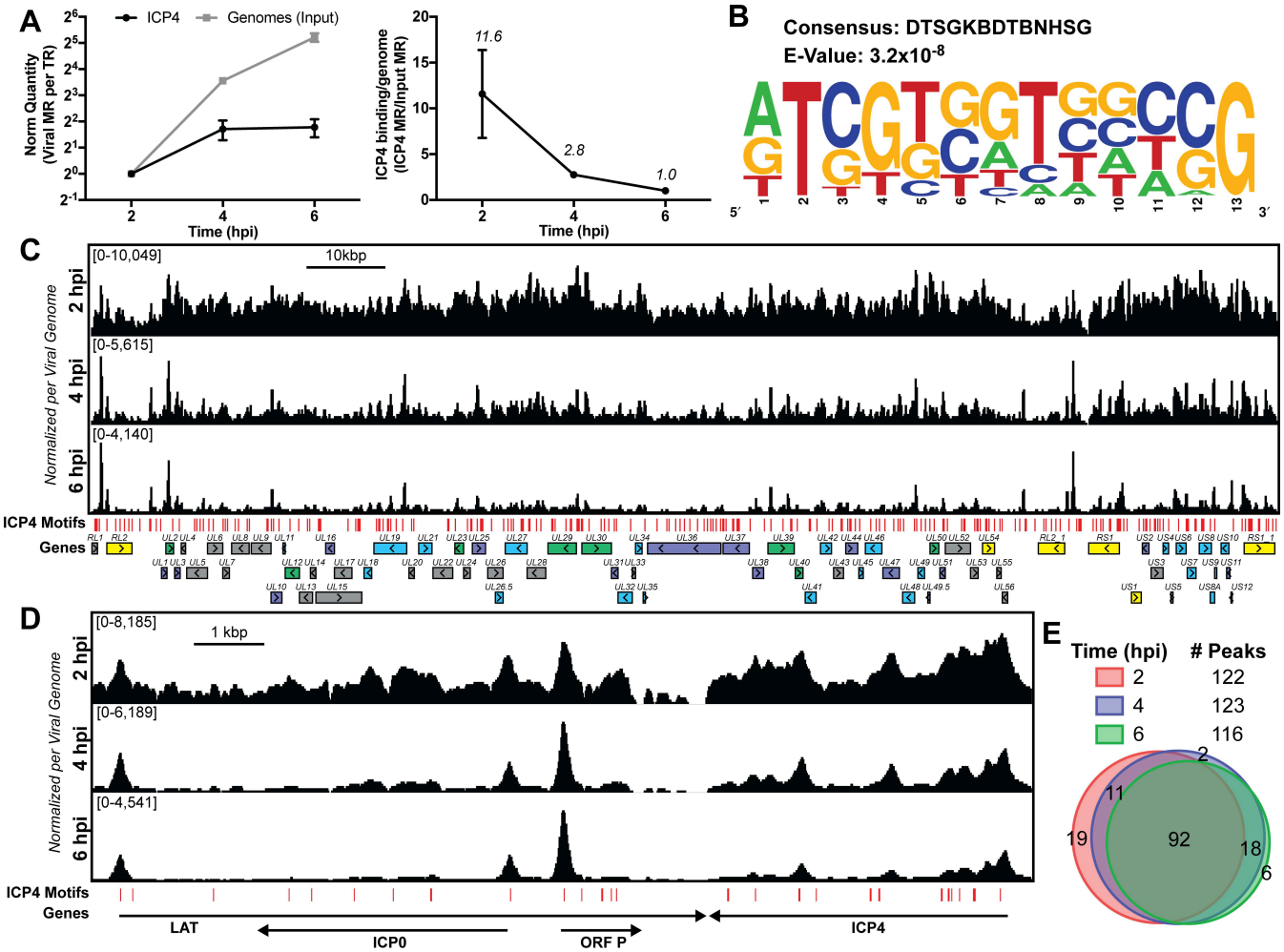
Fig. S6. MRC5 cells were infected with HSV-1 for 2 h, and ChIP-Seq for ICP4, Pol II, H3, H3K4me3, H3K27acetyl, H3K9me3, and H3K27me3 was performed. Data was aligned to the

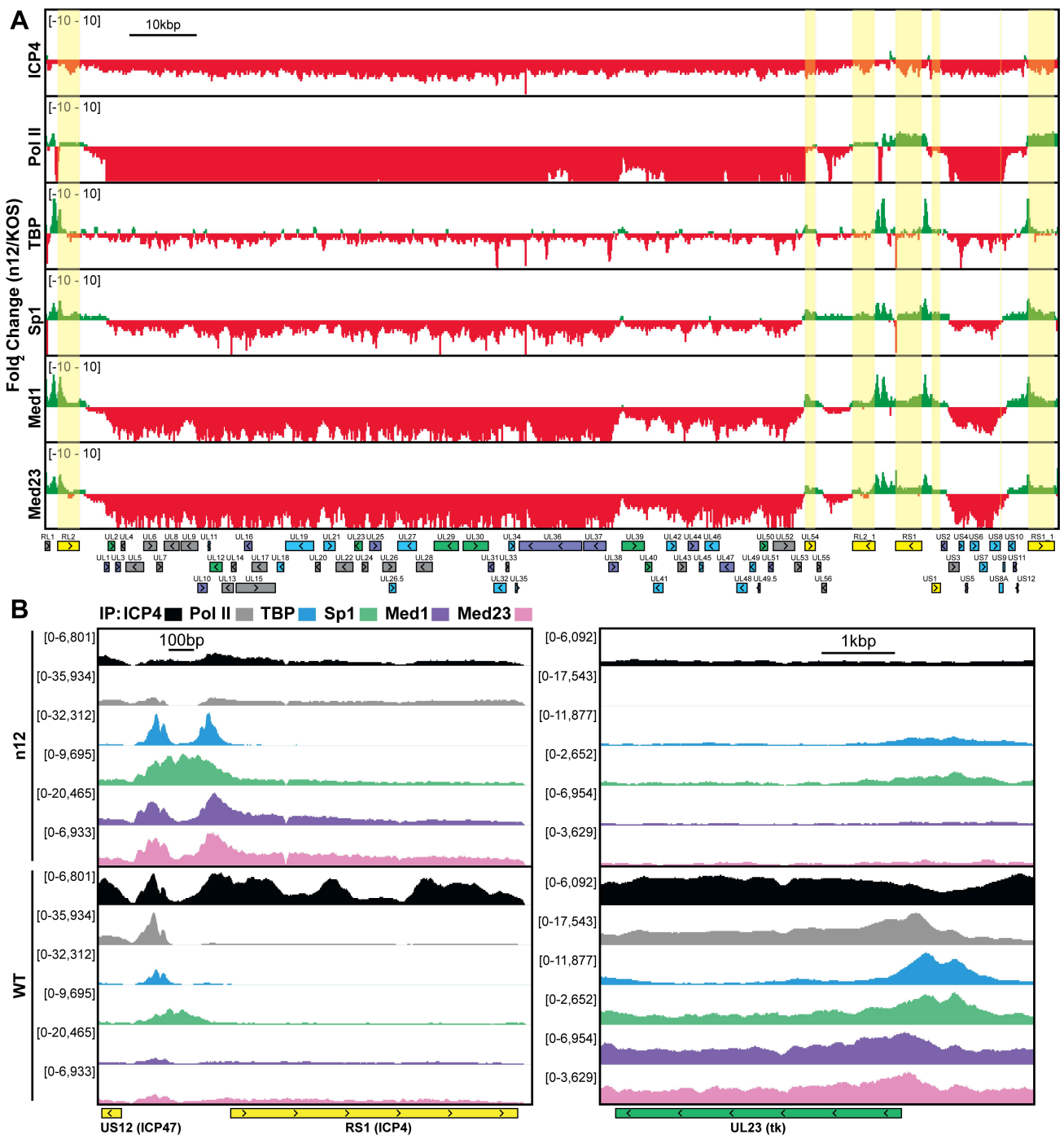
686 human genome (hg38). IP peaks consistent between biological duplicate experiments were
687 determined using MACS2. ChIPSeeker assessment of bound regions for each set of IP
688 peaks.
689

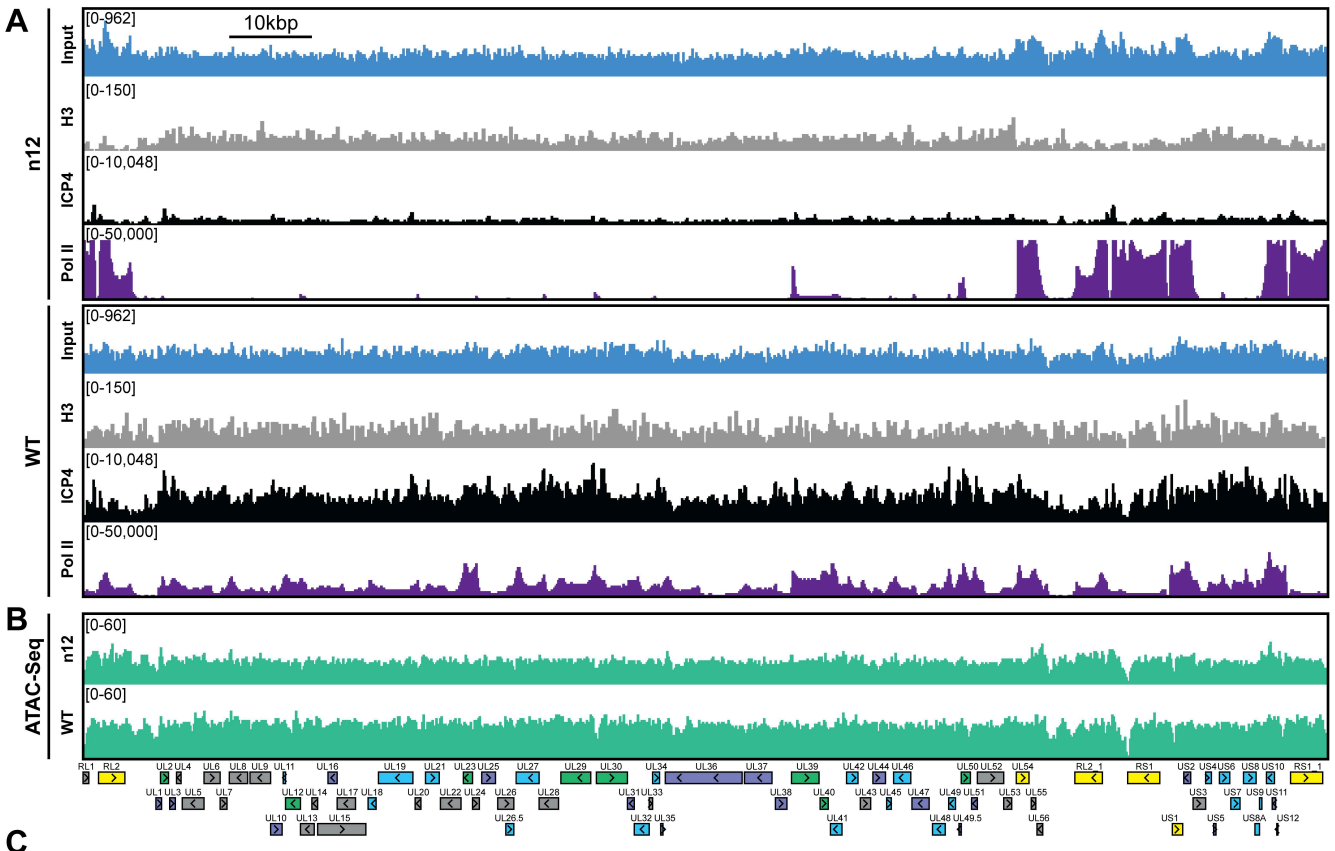
690

691 **Acknowledgements:**

692 This work was supported by NIH grant R01 AI030612 to N.A.D. S.E.D. was supported by the
693 NIH training grants T32AI060525 and F31AI36251. We acknowledge members of the
694 DeLuca lab for thoughtful discussions related to this project and Frances Sivrich for technical
695 assistance.


A
ICP4
Genomes (Input)
Norm Quantity (Viral MR per TR)
Time (hpi)
ICP4 binding/genome (ICP4 MR/Input MR)
11.6
2.8
1.0
B
Consensus: DTSGKBDTBNHSG
E-Value: 3.2x10-8
5′
3′
C
[0-10,049]
10kbp
2 hpi
[0-5,615]
4 hpi
[0-4,140]
6 hpi
Normalized per Viral Genome
ICP4 Motifs
Genes
D
[0-8,185]
1 kbp
2 hpi
[0-6,189]
4 hpi
[0-4,541]
6 hpi
Normalized per Viral Genome
ICP4 Motifs
Genes
LAT
ICP0
ORF P
ICP4
E
Time (hpi)
Peaks
2
122
4
123
6
116
2
11
19
92
18
6


A
Fold Change (n12/KOS)
ICP4
Pol II
TBP
Sp1
Med1
Med23
[-10 - 10]
10kbp
B
IP: ICP4 Pol II TBP Sp1 Med1 Med23
n12
WT
[0-6,801]
[0-35,934]
[0-32,312]
[0-9,695]
[0-20,465]
[0-6,933]
100bp
[0-6,092]
[0-17,543]
[0-11,877]
[0-2,652]
[0-6,954]
[0-3,629]
1kbp
US12 (ICP47)
RS1 (ICP4)
UL23 (tk)

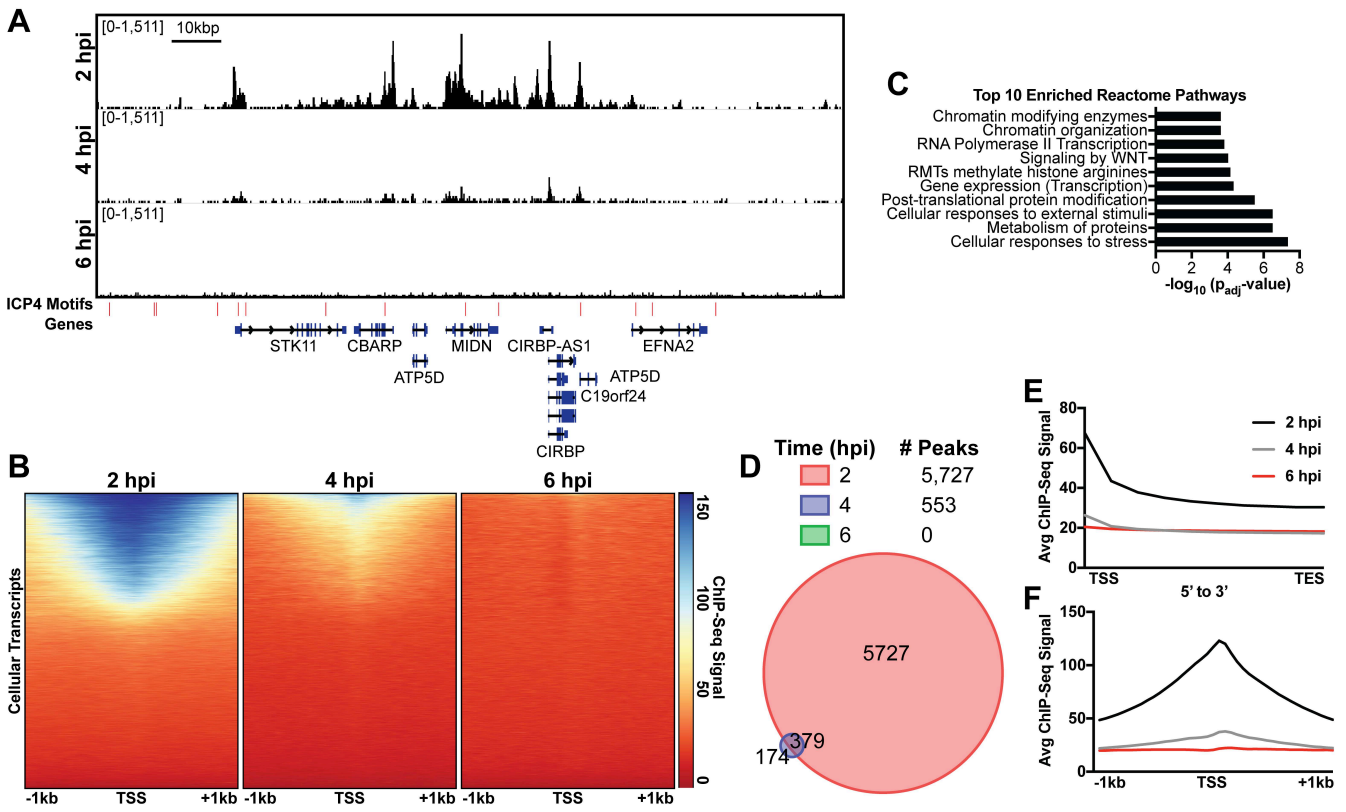


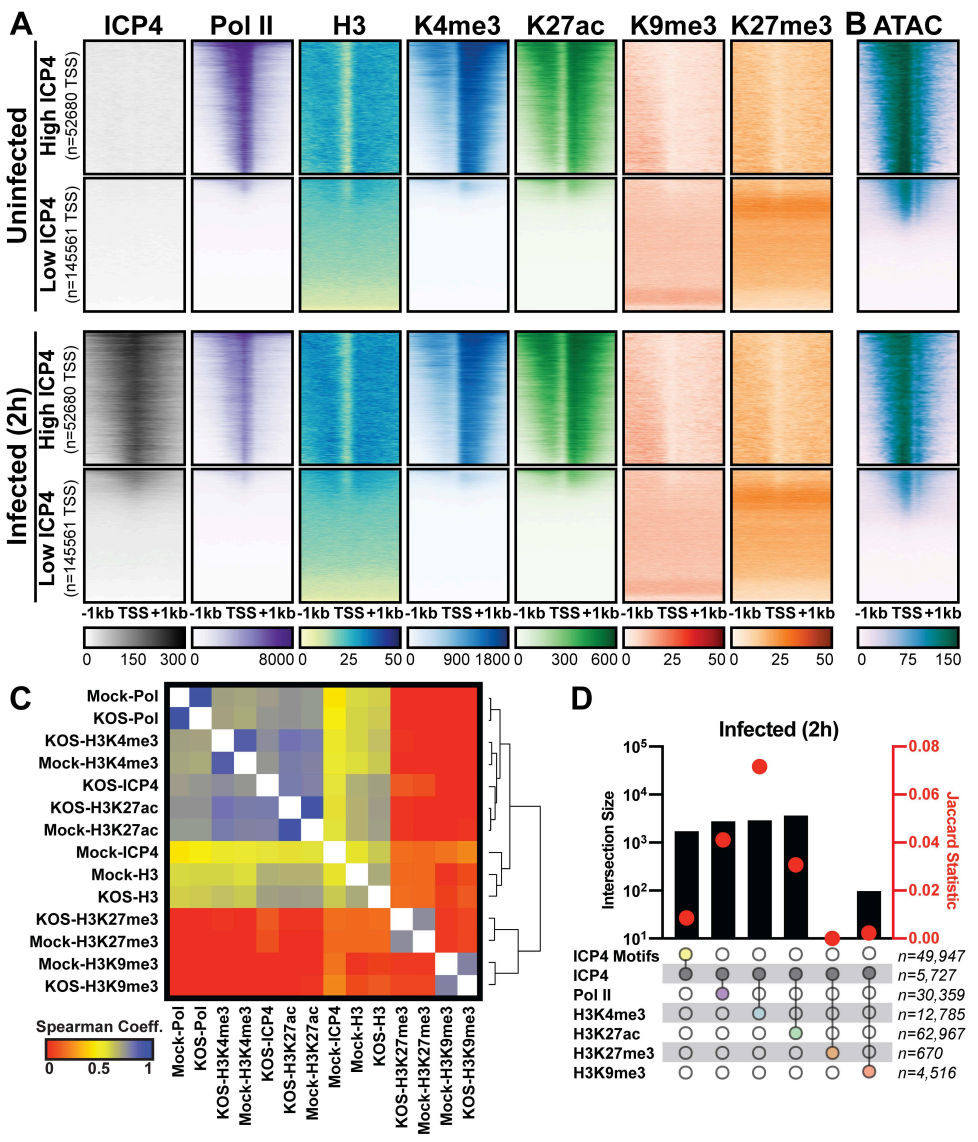
C

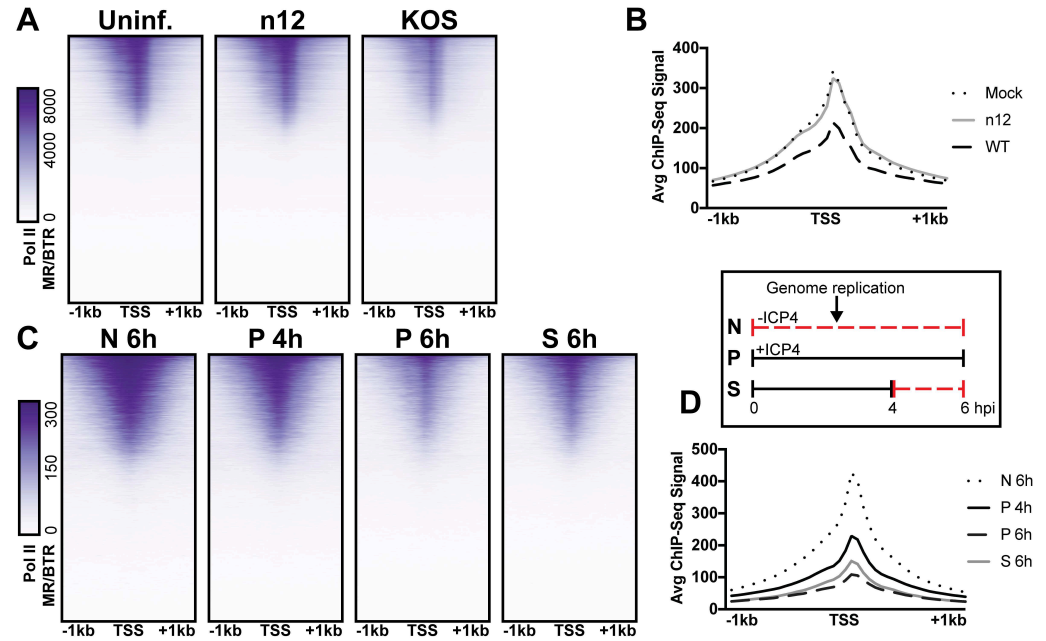
Sample	Genome	Genome Size (bp)	Genome Copy #	Total DNA (bp)	% of Total DNA*	ATAC-Seq Read %	Fold enrichment over expected**
n12	**Host**	3.2E+09	2	6.4E+09	99.4 (±0.4)	98.3	0.99 (±0.03)
	Viral	1.5E+05	254 (±169)	3.9 (±2.6) x10^7	0.6 (±0.4)	1.7	2.8 (±1.1)
	Rel Diff	*2.1E+04*		*377*			
WT 2h	**Host**	3.20E+09	2	6.40E+09	99.6 (±0.01)	98.4	0.99 (±0.0001)
	Viral	1.52E+05	169 (±4)	2.6 (±0.06) x10^7	0.4 (±0.01)	1.6	4.0 (±0.1)
	Rel Diff	*2.13E+04*		*721*			

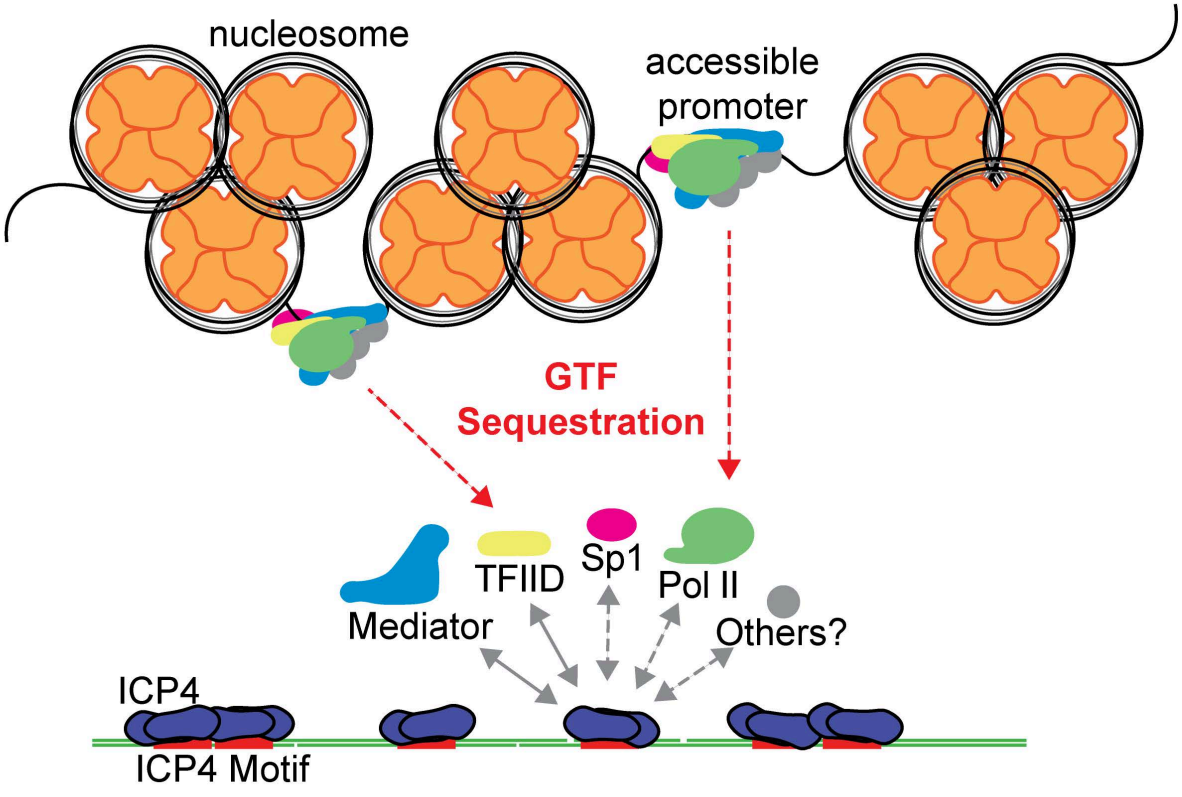
***Based on percent of total reads from ChIP-Seq Input reads, values are the mean (± SD) of four biological replicates**

****Fold enrichment=ATAC-Seq read percentage/percentage of Total DNA**


A
[0-1,511]
10kbp
2 hpi
[0-1,511]
4 hpi
[0-1,511]
6 hpi
ICP4 Motifs
Genes
STK11
CBARP
MIDN
CIRBP-AS1
EFNA2
ATP5D
ATP5D
C19orf24
CIRBP
B
2 hpi
4 hpi
6 hpi
Cellular Transcripts
ChIP-Seq Signal
150
100
50
0
-1kb
TSS
+1kb
C
Top 10 Enriched Reactome Pathways
Chromatin modifying enzymes
Chromatin organization
RNA Polymerase II Transcription
Signaling by WNT
RMTs methylate histone arginines
Gene expression (Transcription)
Post-translational protein modification
Cellular responses to external stimuli
Metabolism of proteins
Cellular responses to stress
0
2
4
6
8
-log10 (padj-value)
D
Time (hpi)
Peaks
2
5,727
4
553
6
0
5727
379
174
E
Avg ChIP-Seq Signal
80
60
40
20
0
2 hpi
4 hpi
6 hpi
TSS
TES
5' to 3'
F
Avg ChIP-Seq Signal
150
100
50
0
-1kb
TSS
+1kb


A
ICP4
Pol II
H3
K4me3
K27ac
K9me3
K27me3
B ATAC
Uninfected
High ICP4 (n=52680 TSS)
Low ICP4 (n=145561 TSS)
Infected (2h)
-1kb TSS +1kb
0 150 300
0 8000
0 25 50
0 900 1800
0 300 600
0 25 50
0 25 50
0 75 150
C
Mock-Pol
KOS-Pol
KOS-H3K4me3
Mock-H3K4me3
KOS-ICP4
KOS-H3K27ac
Mock-H3K27ac
Mock-ICP4
Mock-H3
KOS-H3
KOS-H3K27me3
Mock-H3K27me3
Mock-H3K9me3
KOS-H3K9me3
Spearman Coeff.
0 0.5 1
D
Infected (2h)
Intersection Size
Jaccard Statistic
ICP4 Motifs n=49,947
ICP4 n=5,727
Pol II n=30,359
H3K4me3 n=12,785
H3K27ac n=62,967
H3K27me3 n=670
H3K9me3 n=4,516


A
Uninf.
n12
KOS
Pol II MR/BTR
8000
4000
0
-1kb TSS +1kb
B
Avg ChIP-Seq Signal
400
300
200
100
0
-1kb
TSS
+1kb
Mock
n12
WT
C
N 6h
P 4h
P 6h
S 6h
Pol II MR/BTR
300
150
0
D
Genome replication
N
-ICP4
P
+ICP4
S
0
4
6 hpi
Avg ChIP-Seq Signal
500
400
300
200
100
0
N 6h
P 4h
P 6h
S 6h

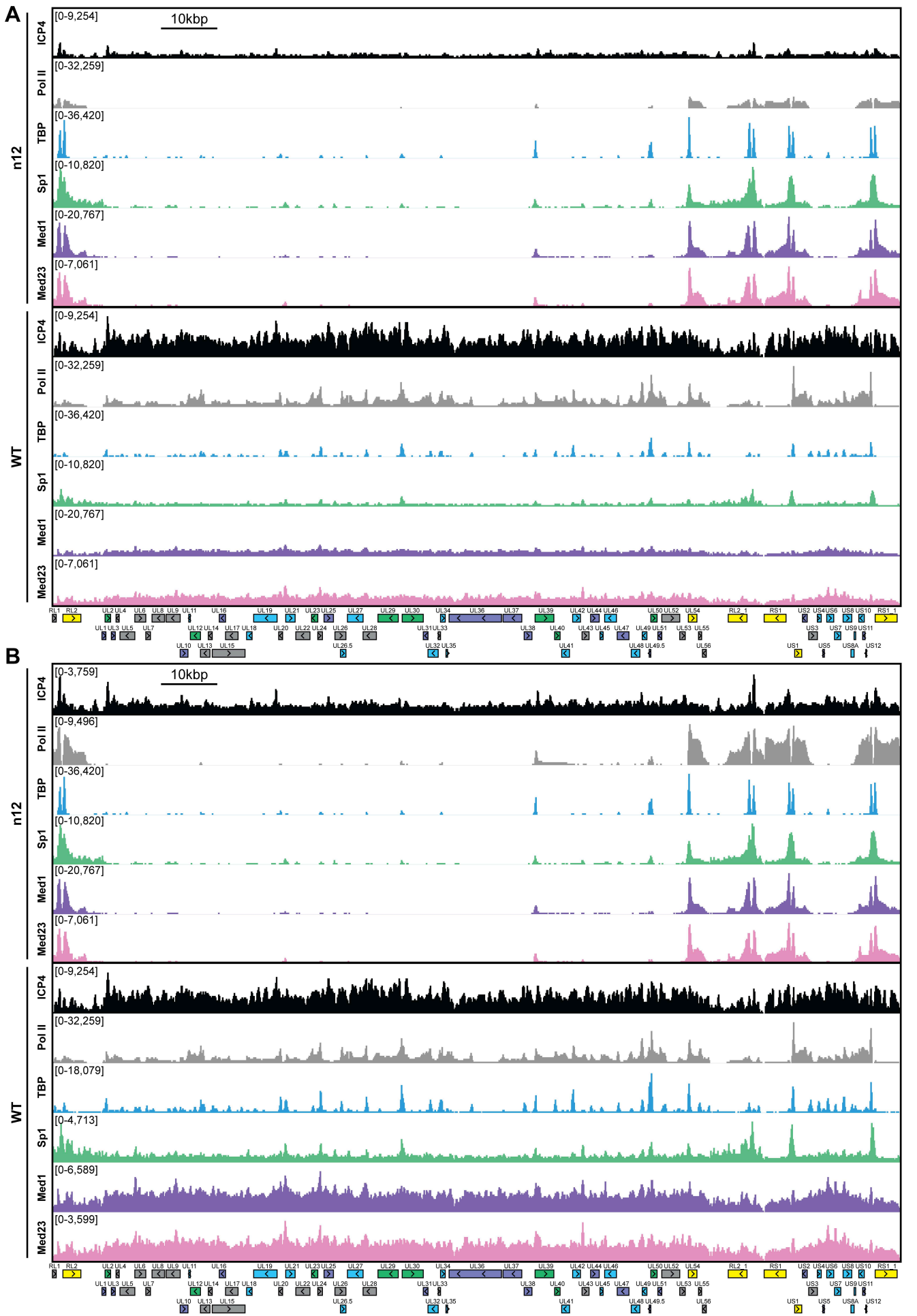


Cellular Genome

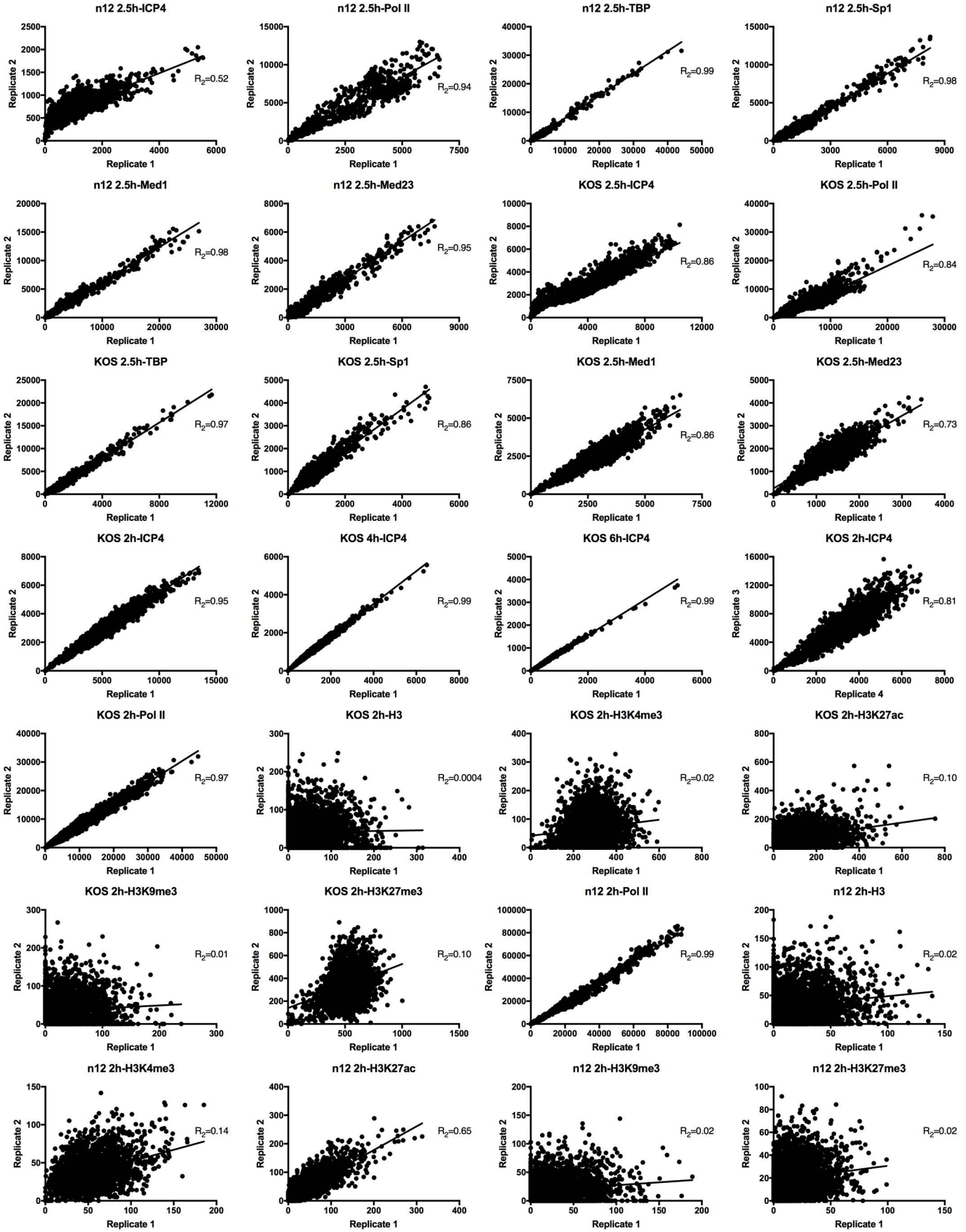
0.02 ICP4 motifs per kbp
0.002 ICP4 peaks per kbp
bound by nucleosomes
minimal Pol II transcription

Viral Genome

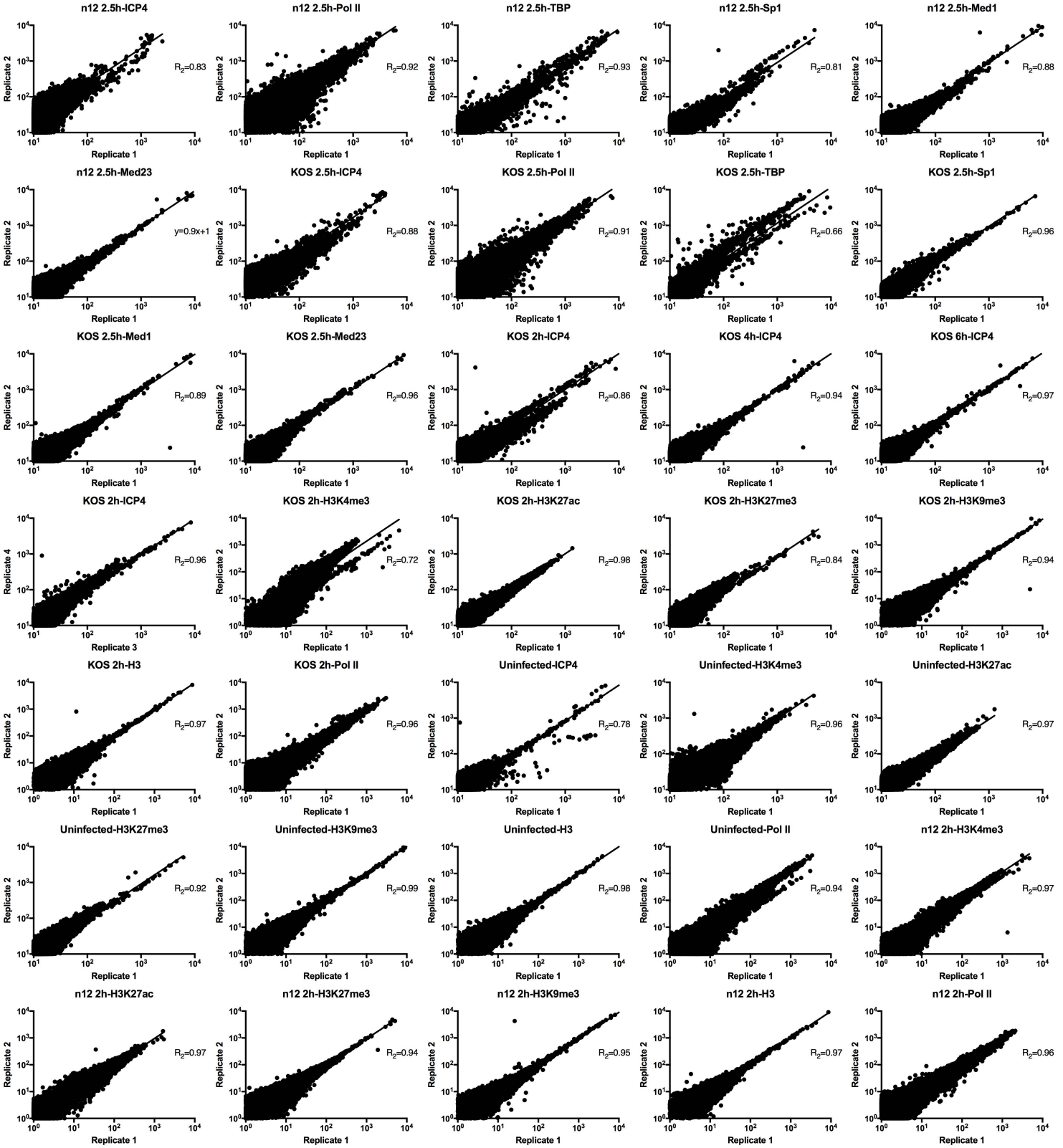
2.0 ICP4 motifs per kbp
0.8 ICP4 peaks per kbp
not associated with nucleosomes
robust Pol II transcription

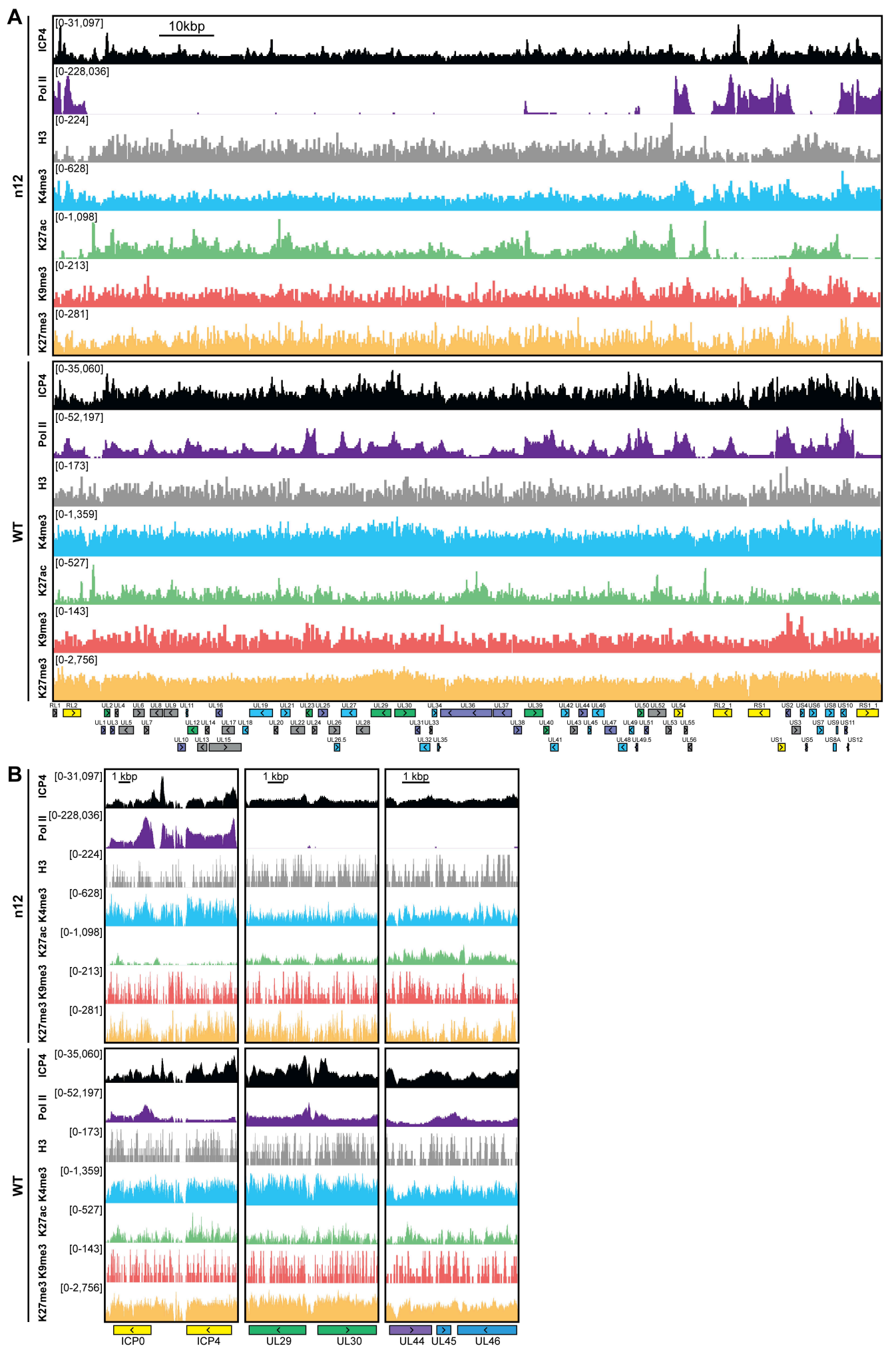

A
n12
WT
ICP4
Pol II
TBP
Sp1
Med1
Med23
[0-9,254]
10kbp
[0-32,259]
[0-36,420]
[0-10,820]
[0-20,767]
[0-7,061]
B
[0-3,759]
[0-9,496]
[0-18,079]
[0-4,713]
[0-6,589]
[0-3,599]

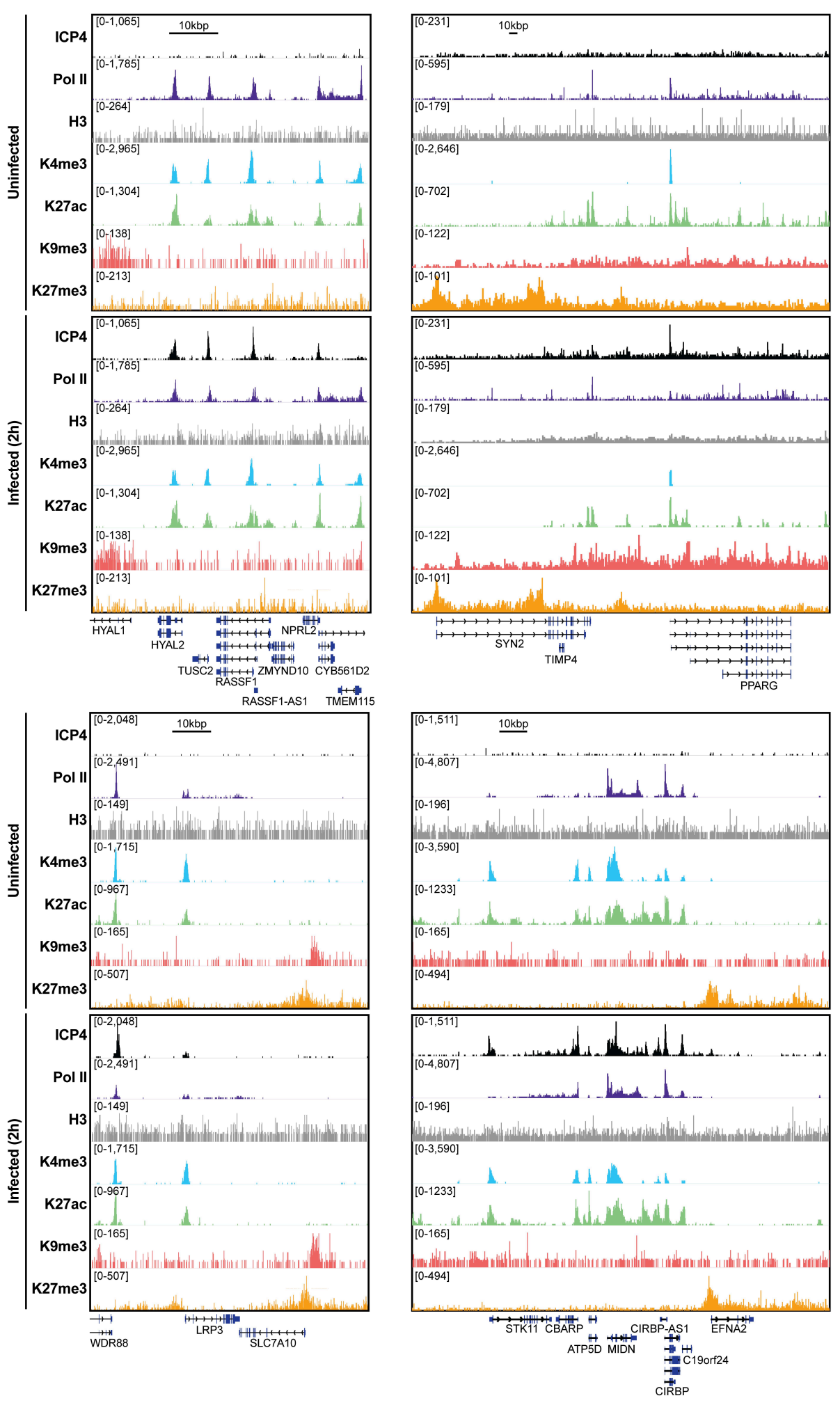
Viral Alignments
bin size=50 bp



Cellular Alignments
bin size=10,000 bp




A
n12
WT
ICP4
Pol II
H3
K4me3
K27ac
K9me3
K27me3
[0-31,097]
10kbp
[0-228,036]
[0-224]
[0-628]
[0-1,098]
[0-213]
[0-281]
[0-35,060]
[0-52,197]
[0-173]
[0-1,359]
[0-527]
[0-143]
[0-2,756]
B
1 kbp
ICP0
ICP4
UL29
UL30
UL44
UL45
UL46


Uninfected
Infected (2h)
ICP4
Pol II
H3
K4me3
K27ac
K9me3
K27me3
10kbp
[0-1,065]
[0-1,785]
[0-264]
[0-2,965]
[0-1,304]
[0-138]
[0-213]
[0-231]
[0-595]
[0-179]
[0-2,646]
[0-702]
[0-122]
[0-101]
HYAL1
HYAL2
NPRL2
TUSC2
RASSF1
ZMYND10
CYB561D2
RASSF1-AS1
TMEM115
SYN2
TIMP4
PPARG
[0-2,048]
[0-2,491]
[0-149]
[0-1,715]
[0-967]
[0-165]
[0-507]
[0-1,511]
[0-4,807]
[0-196]
[0-3,590]
[0-1233]
[0-494]
WDR88
LRP3
SLC7A10
STK11
CBARP
ATP5D
MIDN
CIRBP-AS1
C19orf24
CIRBP
EFNA2