Increased RNA Polymerase Activity and Pausing at CTCF binding sites on the Epstein Barr Virus Genome During Reactivation from Latency

Laura E. M. Dunn[1,2], Fang Lu[3], Paul M. Lieberman[3], Joel D. Baines[1,2*]

[1] Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

[2] Department of Pathobiology, School of Veterinary Medicine, Louisiana State University, Baton Rouge, Louisiana, USA

[3] The Wistar Institute, Philadelphia, Pennsylvania, USA

*Corresponding author

Email: jdb11@cornell.edu

Abstract

The ability of Epstein-Barr Virus (EBV) to switch between latent and lytic infection is key to its long-term persistence, yet the molecular mechanisms behind this switch remain unclear. To investigate transcriptional events during the latent to lytic switch we utilized Precision nuclear Run On followed by deep Sequencing (PRO-Seq) to map cellular RNA polymerase (Pol) activity to single-nucleotide resolution on the host and EBV genome in latently infected Mutu-I cells at 1, 4 and 12 h post-reactivation. During latency, Pol activity was primarily limited to the EBNA1 transcript initiating at the Qp promoter, the EBER and RPMS1/BART regions and the BHLF1 transcript. Unexpectedly at early time-points post-reactivation, the EBV transcripts with the largest increase in Pol activity were LMP-2A, EBER-1 and RPMS1. Closer analysis of the PRO-Seq data at these regions revealed a distinct pattern of high Pol activity with bidirectional transcription and strong peaks indicative of Pol pausing. Alignment to ChIP-Seq data revealed a strong correlation with CTCF binding sites on the EBV genome. In addition, alignment to ATAC-Seq data indicated that many of these transcription regulatory regions were sites of accessible chromatin. Similar features were observed in Akata cells activated from latency with anti-IgG. Overall, these data suggest that during reactivation, EBV recruits RNA polymerase to CTCF binding sites where it transcribes short distances and pauses. These activities likely help open chromatin on the viral genome to initiate productive replication.

Author summary

The ability of EBV to switch between latent and lytic infection is key to its long-term persistence in memory B-cells and its ability to persist in proliferating cells is strongly linked to oncogenesis. During latency, most viral genes are epigenetically silenced,

and the virus must overcome this repression to reactivate lytic replication. Reactivation occurs once the immediate early (IE) EBV lytic genes are expressed. However, the molecular mechanisms behind the switch from the latent transcriptional program to begin transcription of the IE genes remains unknown. In this study, we were able to precisely map RNA polymerase positioning and activity during latency and reactivation. Unexpectedly, RNA polymerase activity was not enriched at the IE genes during early reactivation but accumulated at distinct regions, characteristic of enhancers, on the EBV genome previously shown to be associated with CTCF and open chromatin. We speculate that Pol accumulation at these sites helps maintain open chromatin locally, eventually leading to changes in the quaternary structure of the EBV genome to promote lytic replication.

Introduction

Epstein-Barr virus (EBV) is a gammaherpesvirus that establishes a persistent infection in 95% of adults worldwide. EBV maintains this persistence through a biphasic lifecycle in which it switches between latent and lytic phases. The virus establishes latency in memory B-cells where it exists as a circular episome and the majority of genes are epigenetically silenced via repressive chromatin and DNA methylation (1).

There are four distinct latency types (0-III), which are classified based on the expression of the EBNA proteins, LMP proteins and non-coding RNAs (2). EBNA transcript expression is regulated by promoter selection from the Qp, Cp and Wp latency promoters (2). This promoter selection is mediated by the chromatin architecture of the EBV episome, which is dependent on host CCCTC-binding factor

(CTCF) (3). CTCF is a multifunctional DNA binding protein with known roles in transcriptional activation/repression, chromatin loop and boundary formation and promoter-enhancer blocking activity (4). The EBV genome contains multiple CTCF binding sites (5) and a number of these sites have been linked to the maintenance of the latent chromatin state (6-8).

Reactivation is known to be induced by stimulation of the B cell receptor (BCR) and various stress-related signals (9). These events lead to activation of the viral Zp and Rp promoters to drive transcription of the immediate early (IE) EBV genes, BZLF1 and BRLF1, which in turn induce early lytic gene expression (10). A number of host-cell factors including BLIMP1 (11), PARP1 (8) and MYC/MAX (12) have been shown to be involved in regulating the induction of IE expression. However, the exact molecular mechanisms as to how EBV overcomes the repressive latent state to promote transcription of the IE genes remains unknown.

EBV relies on cellular RNA polymerase II (Pol II) for transcription, although it also uses RNA Polymerase III (Pol III) to transcribe Epstein-Barr virus-encoded small RNA 1 and 2 (EBER1, 2) (13). Pol II contains an extra C-terminal domain (CTD) and the phosphorylation pattern of the CTD is a key factor in the regulation of transcription initiation, elongation and termination (14). EBV is known to modulate Pol II transcription elongation by promoting serine 2 CTD phosphorylation to drive B-cell immortalisation. The virus achieves this by inducing Pol II pausing at the Cp promoter via association with the negative elongation factor (NELF), which facilitates recruitment of the positive transcription elongation factor (pTEFb). pTEFb maintains

CTD serine 2 phosphorylation, leading to strong expression of EBV immortalisation genes (15).

Pausing of Pol II at promoter regions, known as promoter proximal pausing (PPP), is a mechanism known to be used by herpesviruses to regulate transcription. The alphaherpesvirus, Herpes simplex virus-1 (HSV-1) utilizes PPP to regulate temporal gene expression during lytic replication (16). Most notably, Kaposi's sarcoma associated herpesvirus (KSHV), another gammaherpesvirus, exploits PPP to regulate the latent-lytic switch. During latency, NELF is recruited to pause Pol II on lytic KSHV gene promoters allowing for prompt expression upon stimulation of reactivation (17).

The aim of this study was to investigate changes in Pol II activity and positioning during the transition from a highly restricted latent state to the transcriptionally active lytic state. We hypothesized that Pol II remains paused on the promoters of the IE genes during latency and that upon reactivation, Pol is rapidly released into elongation. To address this hypothesis we used Precision nuclear Run On followed by deep Sequencing (PRO-Seq) (18) to map cellular Pol activity to single-nucleotide resolution on the EBV genome in both latently infected Mutu-I cells and Akata cells and at 1, 4 and 12 h post-reactivation.

Using PRO-Seq, we found no evidence of PPP on IE genes during EBV latency or reactivation. Instead, Pol activity and pausing upon reactivation was found to be enriched at distinct regions that were often not associated with any known viral transcript. These regions correlated with CTCF binding sites on the EBV genome,

and the pattern of Pol activity suggested that these CTCF binding sites induced Pol pausing. Alignment to ATAC-Seq data indicated that this stalling of Pol was linked to sites of open chromatin. Overall, this study suggests that one of the first steps during EBV reactivation from latency involves increased bidirectional Pol activity and pausing with an associated opening of chromatin near CTCF binding sites on the viral genome.

Results

Distribution of active polymerase on the EBV episome in latent Mutu-I cells.

In order to assess reactivation, the extent of RNA Pol activity during latency needed to be first established. Nuclei were harvested from the EBV positive Burkitt's lymphoma cell-line, Mutu-I (type 1 latency), and PRO-Seq was performed as previously described (16, 19). Drosophila nuclei were spiked into the run-on reaction and corresponding sequencing reads were used to normalize each sample for sequencing depth. EBV sequences were aligned to the EBV Mutu genome build (20) lacking all but one W repeat to allow appropriate read assignment.

To get an overview of Pol activity on the entire EBV episome during latency, the data was first visualised using the Integrative Genomics Viewer (IGV) (21) (Fig. 1A).To highlight peak regions the IGV windowing function was to set to maximum. Overall, there was a relatively low level of RNA Pol activity across the EBV episome, with the most substantial peak visible at the Pol III-transcribed EBERs and the BHLF1 transcript. To identify the transcripts with the most active Pol, SeqMonk software (22) was used to align and quantify PRO-Seq reads to the viral genome. Most

transcripts aligned with between 10 - 1000 normalised reads (Fig. 1B), indicating a low background level of Pol activity on the viral genome during latency.

The only transcripts with over 1000 reads were BHLF1, RPMS1, both EBERs and multiple EBNA transcripts. The EBERs had the highest level of Pol activity with a mean of 29,061 normalised reads for EBER-1 (Log10 4.46) and 24,062 for EBER-2 (Log10 4.38); this level was over 4-fold higher than any other EBV transcript. As the EBNA genes have overlapping mRNAs we were unable to assign the exact EBNA transcript to reads from this region. However, we were able to quantify reads from the Cp, Qp and Wp latency promoters (Fig. 1C). This analysis revealed that only the Qp promoter, which is known to drive EBNA1 expression, had substantial Pol activity during latency. In summary, PRO-Seq analyses indicated that the highest level of Pol activity in latent Mutu-I cells was on the EBERs and the EBNA transcripts (driven from the Qp promoter). This result is highly consistent with type 1 latency and thus confirmed the validity of this technique for studying EBV transcription.

Distribution and quantification of active polymerase on the EBV episome during reactivation in Mutu-I cells.

We next investigated the transcriptional profile of EBV reactivation using PRO-Seq. Lytic replication in latently infected Mutu-I cells was initiated by treatment with NaB/TPA. Nuclei were harvested at 1, 4 and 12 h post-reactivation, followed by PRO-Seq analysis. Representative whole genome IGV views for all time points are shown in Fig. 2A. To highlight peak regions the IGV windowing function was to set to maximum. Using IGV visualization of the data at the whole genome level, there were only small differences visible overall. However, at 1h post-reactivation, noticeable

peaks emerged at LMP-2A, RPMS1-OriLyt and antisense to the EBERs and these remained at 4 and 12h post reactivation.

To determine whether the changes in Pol activity were significant, the reads for each EBV transcript in latent and reactivated samples were quantified using DeSeq2 Fold change. At 1h post-reactivation, the strongest and most significant increase in Pol activity mapped to LMP-2A, while EBER-1 also had a significant increase (Fig. 2B). Interestingly, most EBV genes had a reduction in Pol activity during this first hour of reactivation. At 4h post-reactivation, Pol activity on LMP-2A remained significantly up-regulated and activity also increased on other LMP genes, LMP-1 and LMP-2B (Fig. 2C). Activity on the RPMS1 transcript also increased significantly at this time point. By 12h post-reactivation, Pol activity was increased on most genes but still remained highest on LMP-2A and RPMS1 (Fig. 2D). The transcriptional activators BZLF1 and BRLF1 are also highlighted in Fig. 2D but significant increases of Pol activity on these genes was only observed starting at 12h post-reactivation.

PRO-Seq identifies distinct patterns of Pol activity on the EBV episome during reactivation.

Because the regions with the strongest increases in Pol activity during reactivation (LMP-2A, RPMS1, EBER) were unexpected, we examined these regions in more detail using IGV. This revealed that the increase in Pol activity on LMP-2A at 1 h post-reactivation was not at the promoter region, but instead was primarily limited to the first exon, located approximately 80 bp downstream of the promoter (Fig. 3A). At 4 and 12 h post-reactivation another peak appeared 150 bp downstream of the initial peak but on the opposite strand, and in a region that does not map to any known

transcript. An additional broad Pol peak appeared on the LMP-2A intronic region at 4 and 12 h post-reactivation, approximately 500 bp downstream of the exon 1 peak. EBER-1 was the only other transcript to have a significant upregulation of Pol activity at 1h post-reactivation. Closer analysis of this region revealed that most of the Pol activity increase was upstream of EBER-1 on both strands, in an area of no known transcripts (Fig. 3B). The RPMS1 region also had an interesting pattern of Pol activity, with Pol activity appearing during reactivation on both strands of the genome, and mapping to introns and regions with no known transcripts (Fig. 3C). Overall, PRO-Seq analysis has indicated that upregulation of Pol activity during EBV reactivation occurs not at gene promoters (as might be expected if promoter proximal pausing was a prominent feature), but more commonly it occurs bidirectionally at intronic or intergenic regions.

dREG detection of transcription regulatory elements from PRO-Seq corresponds with CTCF binding sites on the EBV episome.

The regions with the highest levels of Pol activity during reactivation were unexpected and so we sought to analyze the data in an alternative way. We chose to use the software package dREG (Detection of regulatory elements using GRO-seq and other run-on and sequencing assays), which is trained to detect transcription regulatory elements including transcription initiation regions (TIRs) (23), and applied this to our PRO-Seq data. Novel TIRs discovered by this package have been shown to be associated with transcription factor binding sites. We therefore utilised the EBV portal database (5) to align EBV transcription factor binding ChIP-Seq data to the EBV dREG peaks generated from our PRO-Seq data (Fig. 4A). There were certain regions on the EBV episome where multiple transcription factor binding sites aligned

to the significant dREG pol peaks (highlighted in green, false discovery rate (FDR) ≤ 0.05). For example, CTCF, c-FOS, c-JUN, MAX and RAD21 all had peaks aligning to the EBER dREG peak. CTCF, Sp1, EGR1 and GCN5 also had peaks mapping to dREG peaks at both origins of lytic replication (OriLyt). However, the most striking observation was that of the 8 significant dREG peaks (highlighted in green), a CTCF peak also aligned. Notably, the regions found to be of relevance from DeSeq/PRO-Seq analysis, LMP-2A, EBER and RPMS1, all had aligned dREG/CTCF peaks. With the exception of the Cp peak, all of the significant dREG peaks were also present during latency despite low PRO-Seq reads at most of these regions.

To assess if this was specific to Mutu-I cells and the NaB/TPA method of reactivation we repeated the experiment using another EBV positive Burkitt's lymphoma cell-line, Akata, reactivated with anti-IgG. Reactivation was followed by extraction of nuclei at 1, 4 and 12 hours, and PRO-Seq/dREG analysis was performed as above. It should be noted that EBV copy number varies between the cell types, resulting in the Akata PRO-Seq data having lower read counts (PRO-Seq IGV browser images and volcano plots in Fig. S1). dREG analysis improved the ability to compare peak regions between the cell types. This data showed an association of 7 CTCF peaks with significant dREG pol peaks during latency and reactivation (Fig. 4B) and 6 of these were in common with Mutu-I cells. These were located at; EBERs, OriLyt (left), Qp promoter, RPMS1, OriLyt (right) and LMP-2A (exon 1). All significant dREG peak locations are listed in Table S1. A noticeable difference from Mutu-I was that two significant dREG peaks emerged during reactivation in the middle of the genome near BZLF1 and BKRF3 (overlapped with a CTCF site). In addition, no peaks were apparent close to the Cp and Wp promoters in Akata, as seen in Mutu-I. In

conclusion, dREG analysis has detected a number of significant transcriptional regulatory elements on the EBV genome during latency and reactivation. Though there was a slight variation between the total and location of peaks between Mutu-I and Akata cells, there was a consistent association of dREG peaks with CTCF binding sites in both cell types.

Alignment of CTCF ChIP-Seq and PRO-Seq/dREG indicates pausing of Pol at CTCF sites on EBV genome.

It was noticeable that most dREG/CTCF peak alignments in Fig. 4 were slightly shifted from each other. To further investigate this observation, we used IGV to look at regions of interest at higher resolution using the Mutu-I data. Peaks of Pol accumulation were visible on both strands of the genome next to or over the CTCF peak. Peaks in PRO-Seq are associated with pausing of the polymerase, which leads to an abundance in transcripts adjacent to the pause site (18). This is shown in detail at the LMP-2A region in Fig. 5A, with clear Pol peaks emerged on either side of the CTCF peak during reactivation. A similar pattern was also seen at the RPMS1 CTCF peak, with strong bidirectional Pol activity accumulation around this CTCF peak during reactivation (Fig. 5B). Further examples of bidirectional Pol activity accumulation adjacent to CTCF sites during reactivation are shown for the EBER region (Fig. 5C), at the Qp promoter (Fig. 5D) and the Cp promoter (Fig. 5E). The shifted dREG/CTCF peak pattern was also apparent in Akata cells during reactivation, with these sites also showing strong Pol activity during reactivation (Fig. S2 note: due to the lower EBV genome copy number in Akata, the PRO-Seq signal is weaker and often not visible until 12 h post-reactivation). Overall, these data

suggest that Pol initiates, accumulates, and pauses bidirectionally at regions adjacent to CTCF binding sites during EBV reactivation.

Pol accumulates at sites of accessible chromatin on the EBV episome during early reactivation.

The extension of Pol activity near CTCF binding sites at only 1 h post-reactivation suggests a localized increase in accessible chromatin on the EBV episome. To assess this, we used data from a previous Assay for Transposase-accessible Chromatin Sequencing (ATAC-Seq) experiment (GEO: GSE172476) in 2 sub-populations of latent Mutu-I cells, MUN14 and M14, and aligned this to the Mutu-I dREG and CTCF ChIP-Seq data in IGV (Fig. 6). Five ATAC-Seq peaks aligned strongly with CTCF peaks and with dREG Pol peaks at regions already identified to be of importance, specifically; EBER, Qp, RPMS1, LMP-2A and both OriLyts. This alignment indicates that Pol activity localizes in specific regions of open chromatin during reactivation.

CTCF binding alterations on the EBV episome during reactivation

Finally, we investigated what happens to CTCF binding on the EBV episome during early reactivation. ChIP was performed on latent (untreated) and 4h reactivated (NaB/TPA treated) Mutu-I cells and also latent (untreated) and 4h reactivated (anti-IgG treated) Akata cells using antibodies against CTCF, Pol II and RAD21. qPCR was used to assess binding at multiple CTCF sites on both the EBV and human genome. The Pol II ChIP confirmed an increase in Pol II binding on the EBV genome during reactivation in both Mutu-I and Akata cells, though the reactivation treatment

in both cell types also led to an increase in Pol II at sites on the human genome (Fig. 7A, 7B). At 4h post-reactivation in Mutu-I, there was a decrease in CTCF binding at all 3 sites examined on the EBV genome (LMP-CTCF, Cp-CTCF, Qp-CTCF) (Fig. 7C). In contrast, at 4h post-reactivation in Akata cells, CTCF binding was increased at the LMP-CTCF and Qp-CTCF binding sites and was unchanged at the Cp-CTCF site (Fig. 7D). As CTCF can function alongside the cohesin complex, binding of the cohesin component RAD21 was also assessed. RAD21 binding at the LMP-CTCF locus has already been shown (Fig. 4A) and the ChIP results revealed an increase in RAD21 binding at this site in Mutu-I cells at 4h post-reactivation (Fig. 7E). RAD21 binding at this site was unchanged at 4h post-reactivation in Akata cells (Fig. 7D). Overall, this ChIP study indicates that CTCF binding on the EBV episome is altered during reactivation and could be linked to the cohesin complex. However, the changes are dependent on the cell-line/method of reactivation and the subsequent kinetics of reactivation.

Discussion

In this study we used PRO-Seq to investigate how Pol activity changes on the EBV episome during the switch from the highly restricted latent transcriptional state to productive lytic replication. Despite the well documented activity of BZLF1 as the viral lytic transcriptional activator (24), the molecular mechanisms that initiate transcription of BZLF1 remain unknown. Here we describe important insights as to how Pol activity responds to reactivation signals to aid the transcriptional switch.

We initially speculated that EBV regulates the latent to lytic switch at the elongation step of transcription via release from PPP of Pol II. However, the PRO-Seq data from

EBV latency (Fig. 1A) showed no evidence of PPP on any EBV genes normally expressed during productive infection. Instead, the data documented a highly restricted transcriptional state during latency. Pol activity (above background) was limited to the EBERs, EBNA1 (from Qp promoter) and RPMS1 transcripts (Fig. 1B and 1C), consistent with transcription of these genes in type I latency (25).

At 1 h post-reactivation, peaks of Pol activity did emerge, although none were located at gene promoters. Quantification of the PRO-Seq reads using Deseq2 fold change analysis (26) indicated that the region in Mutu-I cells with the largest increase in Pol activity was the first exon of the latent gene, LMP-2A. In Akata cells, activity at LMP-2A was also strongly increased (Akata volcano plot data shown in Fig. S1) during the first hour of reactivation but due to low EBV read counts in these cells, it was difficult to visualize the precise region of activity increase. EBER-1 and the RPMS1 transcript also had strong increases in Pol activity during early reactivation in Mutu-I cells, and similar to LMP-2A Pol activity at these sites was bidirectional. Bidirectional transcriptional initiation has been shown to be associated with active enhancers on the human genome (27). This feature and recognition by the dREG tool (23) suggest that the increased Pol activity occurs at previously unrecognized transcriptional enhancer regions. The ATAC-Seq data (Fig. 6) also supports the enhancer hypothesis as accessible chromatin is linked to *cis*-regulatory elements, predominantly enhancers and promoters (28).

EBV latency is known to be regulated by enhancer-promoter interactions. Multiple EBV-encoded transcription factors (EBNA-LP, EBNA2, EBNA3C) are associated with *B-cell super-* enhancer regions important for EBV-induced lymphoproliferation

(29). On the EBV genome, *oriP* acts as an enhancer for efficient transcription from the Cp and Wp promoter (30) and *oriLyt* contains an enhancer region required for viral DNA replication (31) that mediates late gene transcription (32). Recently, it has been shown that MYC binds to the *oriLyt* enhancer during latency and MYC depletion leads to reactivation (12). Our data supports the conclusion that enhancer elements at EBV *oriLyt* regions have important regulatory functions in EBV reactivation.

A curious finding was that the IE transcriptional activators, BZLF1 and BRLF1, did not have any significant increase in Pol activity until 12h post-reactivation in Mutu-I cells. This is despite RT-qPCR validation of lytic cycle induction confirming expression of BZLF1 (Fig S3). This could be explained by the enhancer hypothesis as it is possible the PRO-Seq peaks are due to transcription at these potential enhancer regions, leading to the generation of enhancer RNAs (eRNAs) (33). As PRO-Seq measures Pol activity and not stable transcripts, eRNAs, should be detected readily by PRO-Seq. IE gene transcription could then be stimulated through long-range interactions between these enhancer regions and the IE gene promoters (34). Notably, CTCF has been linked to involvement in the activity of enhancer-promoter function (35, 36).

We discovered a strong correlation between Pol activity and CTCF binding sites on the EBV genome. CTCF is known to be important for maintenance of the chromatin state during EBV latency (6, 37) and interestingly, mutation of only the LMP-CTCF locus disrupted the epigenetic state of the latent viral episome (7). We also found that CTCF binding changed at 3 EBV sites upon stimulation of reactivation in Mutu-I

(Fig. 7C). This contrasts with a previous study that indicated CTCF binding is unchanged during EBV reactivation (8). A possible explanation for this difference is that we looked at an earlier time point, 4h, compared to 24h in the above study. Indeed, ChIP-qPCR at 24h post-reactivation in Mutu-I cells showed a recovery in CTCF binding at both the LMP and Cp CTCF loci (Fig. S4). In addition, the cell/virus background and method of reactivation leads to variation in the kinetics of reactivation (38). This could therefore also explain differences between the two studies and the differences we found here between Mutu-I and Akata cells. Our PRO-Seq data indicate that Mutu-I and Akata cells have different reactivation kinetics as Pol activity on IE genes and other lytic genes increased earlier in Akata cells (Fig. S1).

CTCF is involved in maintaining the distinct chromatin loops found in different EBV latency stages (3) and so may also be involved in reactivation by alteration of chromatin conformation. It has already been shown that EBV reactivation involves the recruitment of cellular chromatin remodeling enzymes, mediated by BZLF1 (39). The authors of this study noted an association between BZLF1, CTCF and open-chromatin (ATAC-Seq), supporting a link between CTCF and chromatin modification during EBV reactivation. The effects we observe on CTCF-cohesin DNA binding may reflect changes in EBV chromatin conformation that are likely to contribute to the regulation of EBV reactivation.

The pattern of alignment between CTCF and PRO-Seq peaks was indicative of Pol pausing adjacent to, and at the CTCF sites on the EBV genome. It has recently been shown that CTCF binding sites lead to Pol II stalling on the human genome but that

stalling was independent of CTCF binding (40). Therefore, the association of Pol and CTCF binding sites on the EBV genome during reactivation may not be due to CTCF binding itself. Our ChIP-Seq data did suggest a role for CTCF at certain binding sites, but more studies are needed to confirm whether CTCF loss is functionally important for reactivation. Paused Pol is associated with maintaining DNA accessibility (41) and stalling of Pol has previously been implicated in the prevention of nucleosome assembly on latent EBV promoters (15). This previous work adds support to the hypothesis that Pol accumulation at CTCF binding sites is involved in altering accessibility of the EBV genome during reactivation.

In summary, the PRO-Seq data presented here show that Pol activity during early EBV reactivation occurs at distinct sites on the EBV episome that show similar characteristics to enhancer regions. A strong correlation was identified between CTCF binding sites and Pol activity, and it appeared that these CTCF sites lead to Pol stalling. Our current hypothesis is that Pol activity is initiated at specific enhancer regions during reactivation due to the accessibility of the genome at these sites. Pol accumulates and stalls to aid maintenance of an open chromatin conformation, allowing for increased promoter accessibility eventually leading to productive replication.

Materials and methods

Cells

EBV positive Burkitt's lymphoma cell line Mutu I and Akata were maintained in RPMI 1640 containing 12% FBS and antibiotics (penicillin and streptomycin). Mutu I was

obtained from Dr. Jeffrey Sample, Penn State University Hershey Medical School, PA. Akata was obtained from Dr. Elena Mattia, University of Rome SAPIENZA, Italy.

Drosophila melanogaster S2 (ATCC) cells were grown in Schneider's medium (Lonza) containing 10% foetal bovine serum (FBS) and maintained at 23°C

EBV reactivation

For PRO-seq, Mutu I cells were treated with NaB (1 mM) and TPA (20 ng/ml) for 0, 1, 4, or 12 hrs, and Akata cells were treated with 10 μg/ml of goat anti-human IgG (Sigma I1886) for 0, 1, 4, or 12 hrs. For ChIP assays, Mutu I cells were untreated or treated with NaB (1 mM) and TPA (20 ng/ml) for 4 or 24 hrs, and Akata cells were untreated or treated with 10 μg/ml of goat anti-human IgG for 4 hrs.

Nuclei isolation

Nuclei were isolated following method previously described (42). In brief, cells were washed 2x with ice-cold PBS and then incubated for 10 min on ice with swelling buffer (10mM Tris-HCl [pH 7.5], 10% glycerol, 3mM $CaCl_2$, 2mM, $MgCl2_2$, 0.5mM DTT, protease inhibitors [Roche] and 4 U/ml RNase inhibitors [RNaseOUT, ThermoFisher]). Cells were scraped from plate, pelleted via centrifugation at 600 x g for 10 min (4°C), resuspended in lysis buffer (swelling buffer + 0.5% Igepal) and then incubated on ice for 20 min to release nuclei. Nuclei were pelleted by centrifugation at 1500 x g for 5 min (4°C), washed 2x with lysis buffer with a final wash in storage buffer (50mM Tris-HCl [pH 8.0], 25% glycerol, 5mM MgAcetate, 0.1mM EDTA, 5mM DTT). Nuclei were resuspended in storage buffer, flash frozen in LN_2 and stored at -80°C.

Precision nuclear run-on assay

The nuclear run-on assays were performed as described in (18, 42, 43). Frozen nuclei were thawed on ice and S2 nuclei spiked-in to Mutu-I/Akata cells at a ratio of 1:1000. An equal volume of run-on buffer (10mM Tris-HCl [pH 8.0], 5mM $MgCl_2$, 1mM DTT, 300 mM KCl, 20µM biotin-11-ATP, -UTP, -GTP, -CTP, 0.4 U/ml RNase inhibitor [RNaseOUT ThermoFisher] and 1% Sarkosyl) was added to thawed nuclei. Run-on was performed under constant shaking for 3 min on a vortex shaker at 37°C. Reaction was ended by the addition of TRIzol LS (ThermoFisher) and vortexed to homogenise.

Library preparation

RNA was extracted from run-on nuclei via TRIZOL extraction. RNA underwent base hydrolysis by the addition of 0.2N NaOH and incubation on ice for 20 min. Unincorporated nucleotides were removed using a P-30 column (Bio-Rad). Biotinylated RNA was purified using streptavidin M280 Dynabeads (ThermoFisher) with a series of washes. Bead-bound RNA was washed 2x in ice-cold high salt wash (50mM Tris-HCl [pH 7.4], 2M NaCl, 0.5% Triton X-100, 0.4 U/ml RNase inhibitor [RNaseOUT ThermoFisher]), 2x in ice-cold medium salt wash (10mM Tris-HCl [pH 7.4], 300mM NaCl, 0.1% Triton X-100, 0.4 U/ml RNase inhibitor [RNaseOUT ThermoFisher]) and once in ice-cold low salt wash (5mM Tris-HCl [pH 7.4], 0.1% Triton X-100, 0.4 U/ml RNase inhibitor [RNaseOUT ThermoFisher]). RNA was eluted from beads with two TRIzol extractions.

The 3'-RNA adapter with a 5' phosphate (5'-Phos) and 3' inverted dT (InvdT) (5'-Phos-GAUCGUCGGACUGUAGAACUCUGAAC-3'-InvdT) was ligated to the 3' end of the RNA using T4 RNA ligase I (NEB). RNA was purified as above by binding to streptavidin beads, washing and TRIzol extracted followed by removal of the 5' cap with 10 U of 5'-pyrophosphohydrolase (RppH) (NEB) and 5' end repair with T4 PNK (NEB). The 5'-RNA adapter (5'-CCUUGGCACCCGAGAAUUCCA-3') was ligated to the 5' end of the RNA using T4 RNA ligase I (NEB). RNA was purified as above by binding to streptavidin beads, washing and TRIzol extracted.

RNA was reverse transcribed with SuperScript III reverse transcriptase (ThermoFisher) using the RNA PCR primer 1 (5'-AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA-3'). The cDNA was PCR amplified with Phusion high-fidelity DNA Polymerase (NEB) using barcoded Illumina PCR index primers. Libraries were purified on an 8% polyacrylamide-TBE gel and sequenced on an Illumina NextSeq 500 (GeneLab at Louisiana State University School of Veterinary Medicine).

PRO-Seq analysis

FastQ files were processed using the PRO-Seq pipeline developed by the Danko lab (Cornell) https://github.com/Danko-Lab/utils/tree/master/proseq. First, adapters were trimmed and reads aligned to a concatenated genome file containing hg38, dm3, EBV (Mutu-I or Akata (20)) and rRNA. Note: the EBV genomes were edited to remove all but one W repeat region to aid read alignment. SeqMonk software (22) was used to align the output .bam files (containing the full-length read) to the drosophila genome for normalisation and to the hg38 and EBV genomes for read

quantification. The IGV genome browser (21) was used for visualisation of the output .bw files (containing only the 3' final read position). Fold change analysis was performed using the R package DESeq2 (26).

dREG analysis

The PRO-Seq output .bw files were merged for each replicate using the mergeBigWigs script from the Danko PRO-Seq pipeline https://github.com/Danko-Lab/utils/tree/master/proseq. The merged .bw files were then imported into dREG web-based server, freely available at https://django.dreg.scigap.org/ and dREG analysis performed (23).

ChIP-Seq and ATAC-Seq data

WIG files of ChIP-Seq data from the EBV genome were downloaded from the EBV portal (5), freely available at https://ebv.wistar.upenn.edu/. ATAC-Seq data is available from the NCBI GEO database under accession GSE172476.

Chromatin immunoprecipitation (ChIP)-qPCR Assays.

ChIP-qPCR assays were performed as described previously (44). Quantification of precipitated DNA was determined using real-time PCR and the delta Ct method for relative quantitation (ABI 7900HT Fast Real-Time PCR System). Rabbit IgG (Cell Signaling, 2729S), rabbit anti-CTCF (EMD Millipore, 07-729) and rabbit anti-Rad21 (Abcam, ab992) were used in ChIP assays. Primers for ChIP assays are listed in Supplement Table S2.

Data availability

Raw data and processed BigWig files have been submitted to the GEO database and will be made public upon publication. Peer reviewers can request a secure token to view the data.

Acknowledgements

Portions of this research were conducted with the SuperMic supercomputer maintained at Louisiana State University (http://www.hpc.lsu.edu), and we thank Dr Le Yan for his help in maintaining the PRO-Seq pipeline. We thank Thaya Stoufflet and Dr. Vladimir Chouljenko at the LSU SVM GeneLab for NextSeq 500 sequencing. We also thank Dr. Claire Birkenheuer for her assistance with PRO-Seq technique and interesting discussions.

This work was supported by public health service grants R01 DE017336 and R01 CA093606 to PM and R01 AI141968 and R21 AI 148926 to JDB from the National Institutes of Health.

References

1. Hammerschmidt W. 2015. The Epigenetic Life Cycle of Epstein-Barr Virus. Curr Top Microbiol Immunol 390:103-17.
2. Minarovits J. 2006. Epigenotypes of Latent Herpesvirus Genomes, p 61-80. *In* Doerfler W, Böhm P (ed), DNA Methylation: Development, Genetic Disease and Cancer doi:10.1007/3-540-31181-5_5. Springer Berlin Heidelberg, Berlin, Heidelberg.
3. Tempera I, Klichinsky M, Lieberman PM. 2011. EBV latency types adopt alternative chromatin conformations. PLoS pathogens 7:e1002180-e1002180.
4. Phillips JE, Corces VG. 2009. CTCF: master weaver of the genome. Cell 137:1194-1211.
5. Arvey A, Tempera I, Tsai K, Chen H-S, Tikhmyanova N, Klichinsky M, Leslie C, Lieberman Paul M. 2012. An Atlas of the Epstein-Barr Virus Transcriptome and Epigenome Reveals Host-Virus Regulatory Interactions. Cell Host & Microbe 12:233-245.
6. Tempera I, Wiedmer A, Dheekollu J, Lieberman PM. 2010. CTCF prevents the epigenetic drift of EBV latency promoter Qp. PLoS Pathog 6:e1001048.

7. Chen H-S, Martin KA, Lu F, Lupey LN, Mueller JM, Lieberman PM, Tempera I. 2014. Epigenetic deregulation of the LMP1/LMP2 locus of Epstein-Barr virus by mutation of a single CTCF-cohesin binding site. Journal of virology 88:1703-1713.
8. Lupey-Green LN, Moquin SA, Martin KA, McDevitt SM, Hulse M, Caruso LB, Pomerantz RT, Miranda JL, Tempera I. 2017. PARP1 restricts Epstein Barr Virus lytic reactivation by binding the BZLF1 promoter. Virology 507:220-230.
9. McKenzie J, El-Guindy A. 2015. Epstein-Barr Virus Lytic Cycle Reactivation, p 237-261. *In* Münz C (ed), Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases doi:10.1007/978-3-319-22834-1_8. Springer International Publishing, Cham.
10. Kenney SC, Mertz JE. 2014. Regulation of the latent-lytic switch in Epstein-Barr virus. Seminars in cancer biology 26:60-68.
11. Reusch JA, Nawandar DM, Wright KL, Kenney SC, Mertz JE. 2015. Cellular differentiation regulator BLIMP1 induces Epstein-Barr virus lytic reactivation in epithelial and B cells by activating transcription from both the R and Z promoters. Journal of virology 89:1731-1743.
12. Guo R, Jiang C, Zhang Y, Govande A, Trudeau SJ, Chen F, Fry CJ, Puri R, Wolinsky E, Schineller M, Frost TC, Gebre M, Zhao B, Giulino-Roth L, Doench JG, Teng M, Gewurz BE. 2020. MYC Controls the Epstein-Barr Virus Lytic Switch. Molecular Cell 78:653-669.e8.
13. Arrand JR, Rymo L. 1982. Characterization of the major Epstein-Barr virus-specific RNA in Burkitt lymphoma-derived cells. J Virol 41:376-89.
14. Buratowski S. 2009. Progression through the RNA Polymerase II CTD Cycle. Molecular Cell 36:541-546.
15. Palermo RD, Webb HM, West MJ. 2011. RNA polymerase II stalling promotes nucleosome occlusion and pTEFb recruitment to drive immortalization by Epstein-Barr virus. PLoS pathogens 7:e1002334-e1002334.
16. Birkenheuer CH, Baines JD. 2020. RNA Polymerase II Promoter-Proximal Pausing and Release to Elongation Are Key Steps Regulating Herpes Simplex Virus 1 Transcription. Journal of virology 94:e02035-19.
17. Toth Z, Brulois KF, Wong L-Y, Lee H-R, Chung B, Jung JU. 2012. Negative elongation factor-mediated suppression of RNA polymerase II elongation of Kaposi's sarcoma-associated herpesvirus lytic gene expression. Journal of virology 86:9696-9707.
18. Kwak H, Fuda NJ, Core LJ, Lis JT. 2013. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science 339:950-3.
19. Birkenheuer CH, Danko CG, Baines JD, Sandri-Goldin RM. 2018. Herpes Simplex Virus 1 Dramatically Alters Loading and Positioning of RNA Polymerase II on Host Genes Early in Infection. Journal of Virology 92:e02184-17.
20. Lin Z, Wang X, Strong MJ, Concha M, Baddoo M, Xu G, Baribault C, Fewell C, Hulme W, Hedges D, Taylor CM, Flemington EK. 2013. Whole-genome sequencing of the Akata and Mutu Epstein-Barr virus strains. J Virol 87:1172-82.
21. Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178-92.
22. Institute B. SeqMonk. http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/. Accessed

23. Wang Z, Chu T, Choate LA, Danko CG. 2019. Identification of regulatory elements from nascent transcription using dREG. Genome Res 29:293-303.
24. Murata T. 2014. Regulation of Epstein–Barr virus reactivation from latency. Microbiology and Immunology 58:307-317.
25. Price AM, Luftig MA. 2015. To Be or Not IIb: A Multi-Step Process for Epstein-Barr Virus Latency Establishment and Consequences for B Cell Tumorigenesis. PLOS Pathogens 11:e1004656.
26. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.
27. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jørgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Baillie JK, Ishizu Y, Shimizu Y, Furuhata E, Maeda S, Negishi Y, Mungall CJ, Meehan TF, Lassmann T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Müller F, Forrest ARR, Carninci P, Rehli M, Sandelin A, et al. 2014. An atlas of active enhancers across human cell types and tissues. Nature 507:455-461.
28. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, Giste E, Johnson AK, Johnson EM, Kutyavin T, Lajoie B, Lee B-K, Lee K, London D, Lotakis D, Neph S, Neri F, Nguyen ED, Qu H, Reynolds AP, Roach V, Safi A, Sanchez ME, Sanyal A, Shafer A, Simon JM, Song L, Vong S, Weaver M, Yan Y, Zhang Z, Zhang Z, Lenhard B, Tewari M, Dorschner MO, Hansen RS, et al. 2012. The accessible chromatin landscape of the human genome. Nature 489:75-82.
29. Zhou H, Schmidt SCS, Jiang S, Willox B, Bernhardt K, Liang J, Johannsen EC, Kharchenko P, Gewurz BE, Kieff E, Zhao B. 2015. Epstein-Barr virus oncoprotein super-enhancers control B cell growth. Cell host & microbe 17:205-216.
30. Puglielli MT, Woisetschlaeger M, Speck SH. 1996. oriP is essential for EBNA gene promoter activity in Epstein-Barr virus-immortalized lymphoblastoid cell lines. J Virol 70:5758-68.
31. Hammerschmidt W, Sugden B. 1988. Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus. Cell 55:427-33.
32. Djavadian R, Chiu Y-F, Johannsen E. 2016. An Epstein-Barr Virus-Encoded Protein Complex Requires an Origin of Lytic Replication In Cis to Mediate Late Gene Transcription. PLoS pathogens 12:e1005718-e1005718.
33. Lam MTY, Li W, Rosenfeld MG, Glass CK. 2014. Enhancer RNAs and regulated transcriptional programs. Trends in Biochemical Sciences 39:170-182.
34. Higgs DR. 2020. Enhancer–promoter interactions and transcription. Nature Genetics 52:470-471.
35. Kubo N, Ishii H, Xiong X, Bianco S, Meitinger F, Hu R, Hocker JD, Conte M, Gorkin D, Yu M, Li B, Dixon JR, Hu M, Nicodemi M, Zhao H, Ren B. 2021. Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. Nat Struct Mol Biol 28:152-161.
36. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y, Lu Y, Wu Y, Jia Z, Li W, Zhang MQ, Ren B, Krainer AR, Maniatis T, Wu Q.

2015. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell 162:900-10.
37. Lupey-Green LN, Caruso LB, Madzo J, Martin KA, Tan Y, Hulse M, Tempera I, Jung JU. 2018. PARP1 Stabilizes CTCF Binding and Chromatin Structure To Maintain Epstein-Barr Virus Latency Type. Journal of Virology 92:e00755-18.
38. Countryman J, Gradoville L, Bhaduri-McIntosh S, Ye J, Heston L, Himmelfarb S, Shedd D, Miller G. 2009. Stimulus duration and response time independently influence the kinetics of lytic cycle reactivation of Epstein-Barr virus. Journal of virology 83:10694-10709.
39. Schaeffner M, Mrozek-Gorska P, Buschle A, Woellmer A, Tagawa T, Cernilogar FM, Schotta G, Krietenstein N, Lieleg C, Korber P, Hammerschmidt W. 2019. BZLF1 interacts with chromatin remodelers promoting escape from latent infections with EBV. Life Science Alliance 2:e201800108.
40. Luan J, Xiang G, Gómez-García PA, Tome JM, Zhang Z, Vermunt MW, Zhang H, Huang A, Keller CA, Giardine BM, Zhang Y, Lan Y, Lis JT, Lakadamyali M, Hardison RC, Blobel GA. 2021. Distinct properties and functions of CTCF revealed by a rapidly inducible degron system. Cell reports 34:108783-108783.
41. Jonkers I, Kwak H, Lis JT. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. Elife 3:e02407.
42. Birkenheuer CH, Danko CG, Baines JD. 2018. Herpes Simplex Virus 1 Dramatically Alters Loading and Positioning of RNA Polymerase II on Host Genes Early in Infection. Journal of virology 92:e02184-17.
43. Mahat DB, Kwak H, Booth GT, Jonkers IH, Danko CG, Patel RK, Waters CT, Munson K, Core LJ, Lis JT. 2016. Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat Protoc 11:1455-76.
44. Chau Charles M, Lieberman Paul M. 2004. Dynamic Chromatin Boundaries Delineate a Latency Control Region of Epstein-Barr Virus. Journal of Virology 78:12308-12319.

Figure Legends

Figure 1: Distribution of active polymerase on the EBV episome in latent Mutu-I cells. (A) Genome browser view of the distribution of PRO-Seq reads (normalised to *Drosophila melanogaster* spike-in) along the EBV episome in latent Mutu-I cells. Representative sample of 2-3 biological replicates. The W repeat region was deleted to include only 1 repeat for the sequencing alignment. The IGV window was set to maximum for each track. (B) Violin plot of the Log10 transformed mean normalised read count of each EBV transcript during latency. The transcripts with a mean of

>1000 normalised reads are labelled. (C) Normalised sequencing read count of each EBV latency promoter during latency in Mutu-I cells. Data represents mean of 3 biological replicates, error bars indicate SEMs.

Figure 2: Distribution and quantification of active polymerase on the EBV episome during reactivation in Mutu-I cells. (A) Genome browser view of the distribution of PRO-Seq reads (normalised to *Drosophila melanogaster* spike-in) along the EBV episome in latent Mutu-I cells and at 1, 4 and 12h post-reactivation with NaB/TPA. Representative sample of 2-3 biological replicates. The W repeat region was deleted to include only 1 repeat for sequencing alignment purposes. The IGV window was set to maximum for each track. Volcano plots showing the DeSeq2 calculated fold changes of PRO-Seq reads for each EBV transcript relative to latency at (B) 1h post-reactivation, (C) 4h post-reactivation and (D) 12h post-reactivation.

Figure 3: PRO-Seq identifies distinct patterns of Pol activity on the EBV episome during reactivation. High resolution IGV view of PRO-Seq tracks from EBV genome regions with the most significant increase in Pol activity during early reactivation in Mutu-I cells. Cells were treated with NaB/TPA to induce reactivation. (A) LMP-2A, exon 1 region. (B) EBER region. (C) RPMS1 region. Representative sample of 2-3 biological replicates. Sequencing reads were normalised to *Drosophila melanogaster* spike-in. PRO-Seq peaks of interest are highlighted in purple. Solid red indicates coding regions, black lines indicate non-coding regions of transcripts. Transcription start sites and promoter regions are in green.

Figure 4: dREG detection of transcription regulatory elements from PRO-Seq corresponds with CTCF binding sites on the EBV episome. (A) IGV view of the PRO-Seq results from latent and reactivated (NaB/TPA treated) Mutu-I cells after analysis with the dREG software package to identify transcription initiation regions (TIRs), aligned to transcription factor ChIP-Seq on the latent EBV episome. (B) IGV view of PRO-Seq results from latent and reactivated (anti-IgG treated) Akata cells after analysis with dREG aligned to CTCF ChIP-Seq on the EBV episome. Significant dREG peaks are highlighted in green (false discovery rate (FDR) ≤0.05). ChIP-Seq data from (5).

Figure 5: Alignment of CTCF ChIP-Seq peaks, dREG and PRO-Seq indicates pausing of Pol at CTCF sites on EBV genome. IGV view of aligned dREG, PRO-Seq and CTCF ChIP-Seq peaks from latent and reactivated EBV in Mutu-I cells (TPA/NaB). (A) LMP-2A, exon 1 region. (B) RPMS1 region. (C) EBER region. (D) Qp promoter. (E) Cp promoter. Representative sample of 2-3 biological replicates. Significant dREG peaks are highlighted in green (false discovery rate (FDR) ≤0.05).

Figure 6: Pol accumulates at sites of accessible chromatin on the EBV episome during early reactivation. IGV view of the PRO-Seq results from latent and reactivated (NaB/TPA treated) Mutu-I cells after analysis with the dREG software package to identify transcription initiation regions (TIRs). Aligned in comparison to CTCF ChIP-Seq and ATAC-Seq from 2 sub populations of latent Mutu-I cells on the EBV episome. Significant dREG peaks are highlighted in green (false discovery rate (FDR) ≤0.05).

Figure 7: CTCF binding alterations on the EBV episome during reactivation. ChIP-qPCR at CTCF binding sites on the EBV and human genomes in latent and at 4h post-reactivation Mutu-I cells (NaB/TPA treated) (A, C and E) and latent 4h post-reactivation Akata cells (anti-IgG treated) (B, D and F). Using antibodies against (A, B) Pol II, (B, C) CTCF and (E, F) RAD21.

Supporting Information

Figure S1: Distribution and quantification of active polymerase on the EBV episome during reactivation in Akata cells. (A) Genome browser view of the distribution of PRO-Seq reads (normalised to *Drosophila melanogaster* spike-in) along the EBV episome in latent Akata cells and at 1, 4 and 12h post-reactivation with anti-IgG. Representative sample of 3 biological replicates. The W repeat region was deleted to include only 1 repeat for sequencing alignment purposes. The IGV window was set to maximum for each track. Volcano plots showing the DeSeq2 calculated fold changes of PRO-Seq reads for each EBV transcript relative to latency at (B) 1h post-reactivation, (C) 4h post-reactivation and (D) 12h post-reactivation.

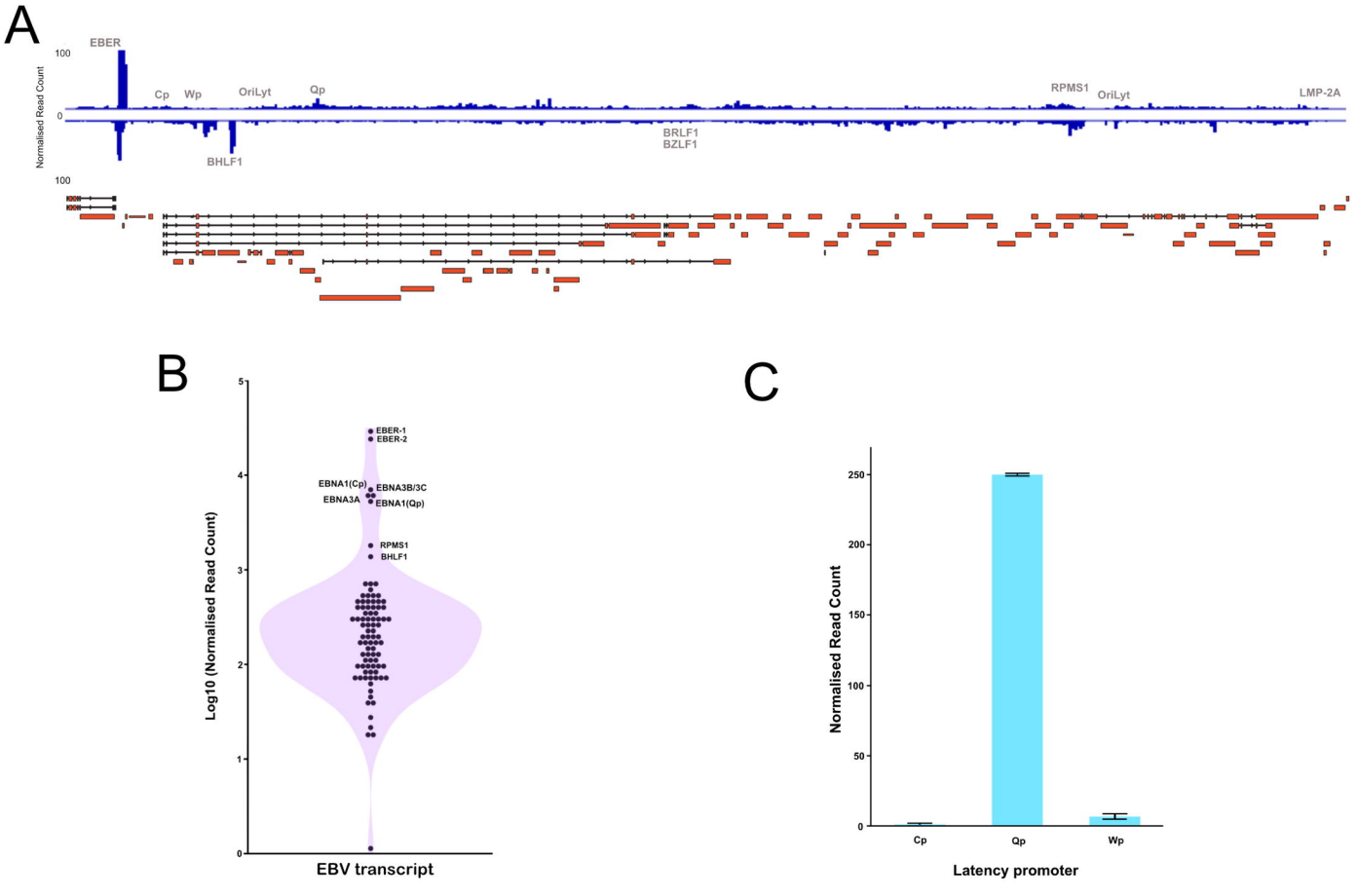
Figure S2: Alignment of CTCF ChIP-Seq peaks, dREG and PRO-Seq from Akata cells indicates pausing of Pol at CTCF sites on EBV genome. IGV view of aligned dREG, PRO-Seq and CTCF ChIP-Seq peaks from latent and reactivated EBV in Akata cells (Anti-IgG treated). (A) LMP-2A, exon 1 region. (B) RPMS1 region. (C) EBER region. (D) Qp promoter. (E) BKRF region. Representative sample of 3 biological replicates. Significant peaks are highlighted in green (false discovery rate (FDR) ≤0.05).

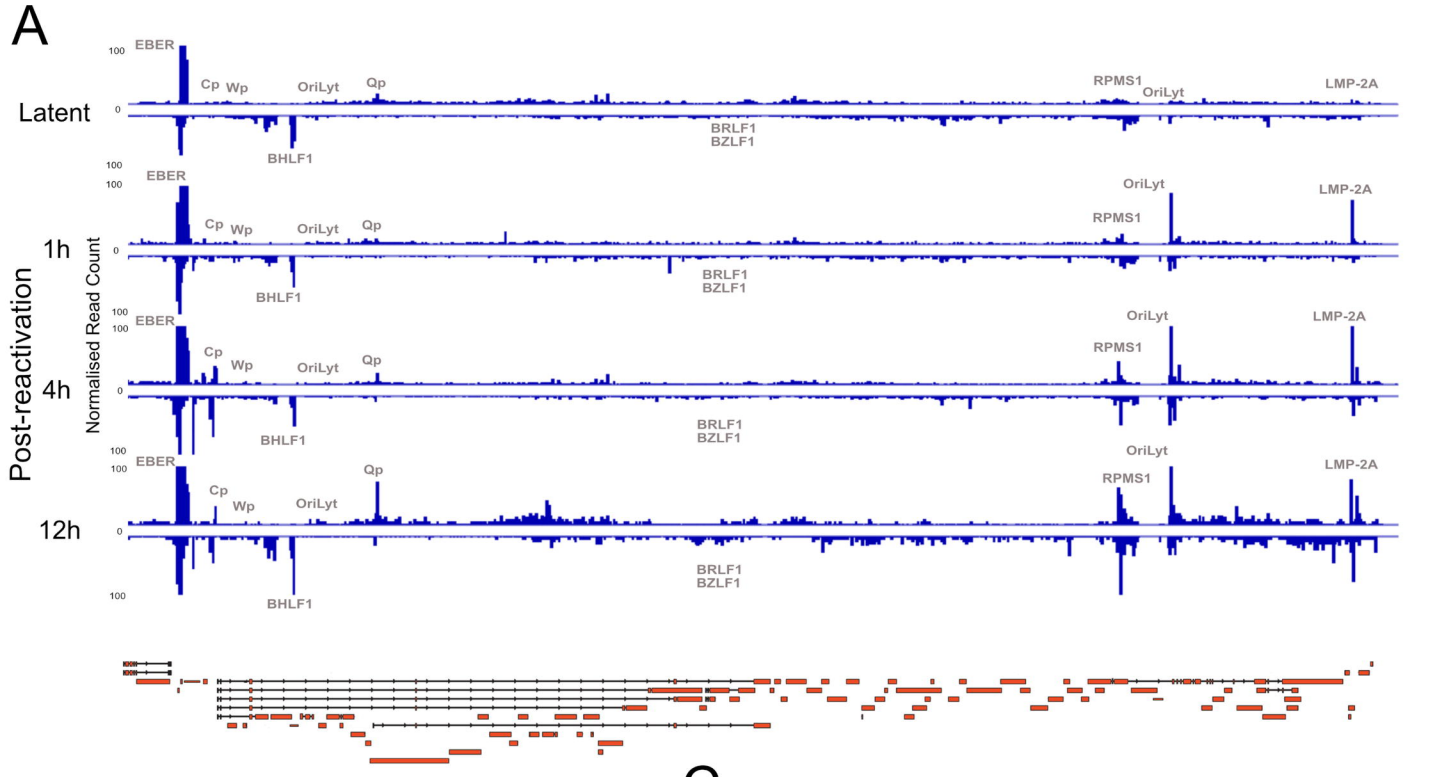
Figure S3: RT-qPCR validation of EBV reactivation. Relative expression of EBV transcripts in untreated and reactivated cells. (A) Mutu-I cells untreated and 24 h post-treatment with NaB/TPA. (B) Akata cells untreated and at 4 and 24 h post-treatment with anti-IgG. Zta: BZLF1, RTA: BRLF1, EA-D: early antigen D.

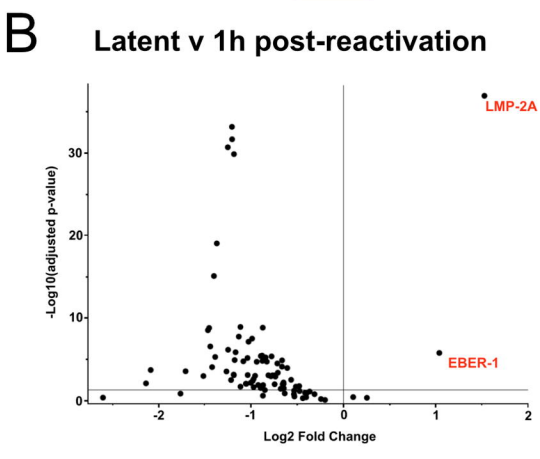
Figure S4: CTCF binding alterations on the EBV episome during reactivation. ChIP-qPCR at CTCF binding sites on the EBV and human genomes in latent and at 4 and 24 h post-reactivation Mutu-I cells (NaB/TPA treated) using antibody against CTCF.

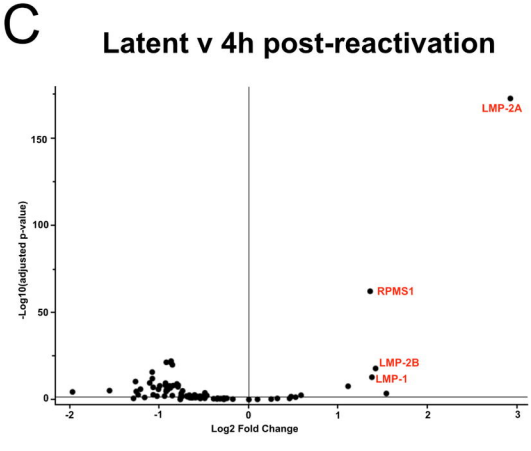
Table S1: Location of dREG peaks on the NC_007605.1 EBV genome in Mutu-I and Akata cells

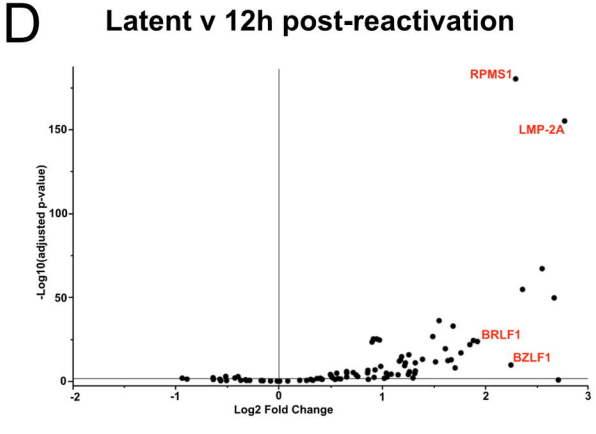
Table S2: Primers used for ChiP-qPCR

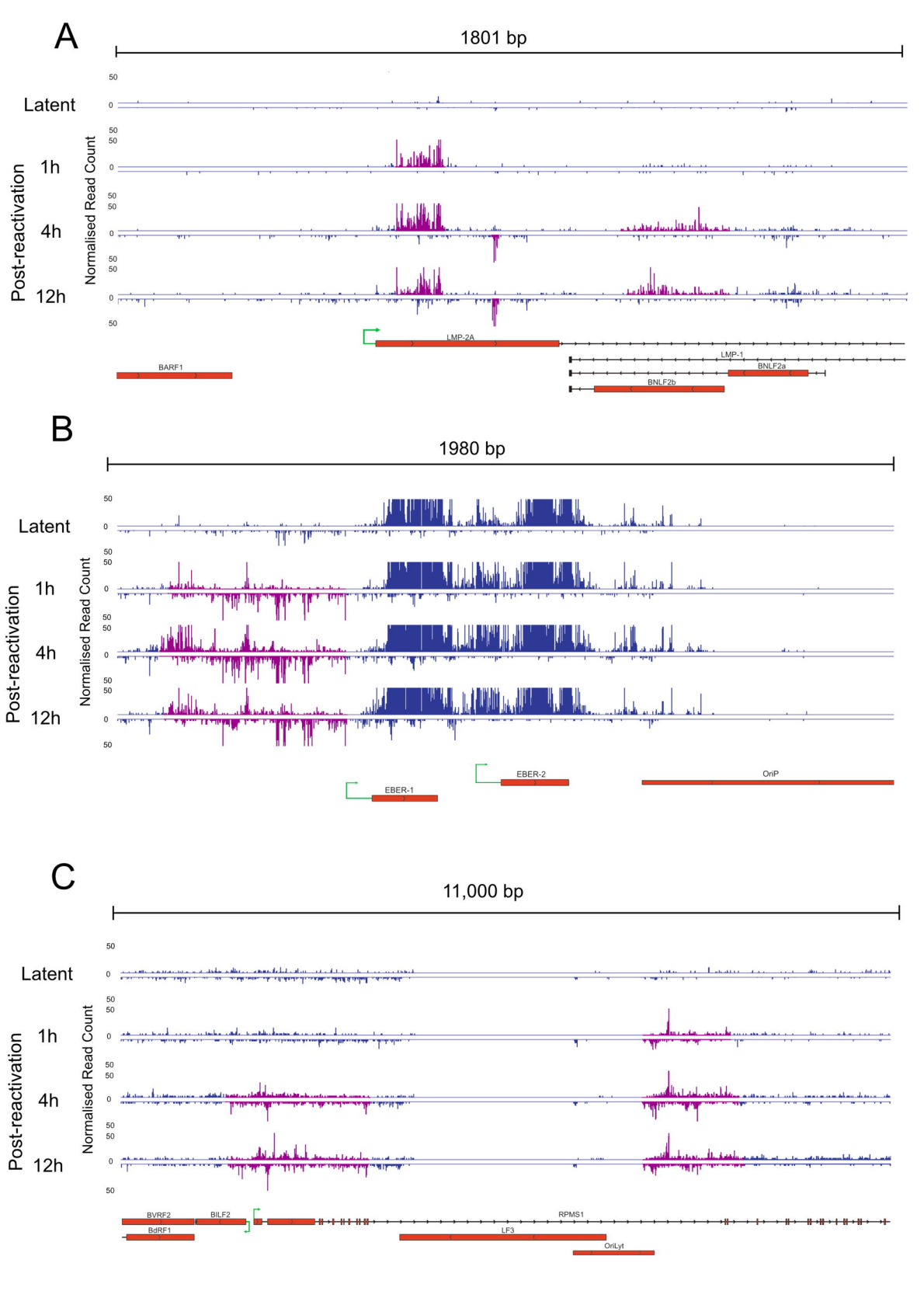

A
EBER
100
Normalised Read Count
0
Cp
Wp
OriLyt
Qp
BHLF1
BRLF1
BZLF1
RPMS1
OriLyt
LMP-2A
100
B
EBER-1
EBER-2
EBNA1(Cp)
EBNA3B/3C
EBNA3A
EBNA1(Qp)
RPMS1
BHLF1
Log10 (Normalised Read Count)
EBV transcript
C
Normalised Read Count
Cp
Qp
Wp
Latency promoter

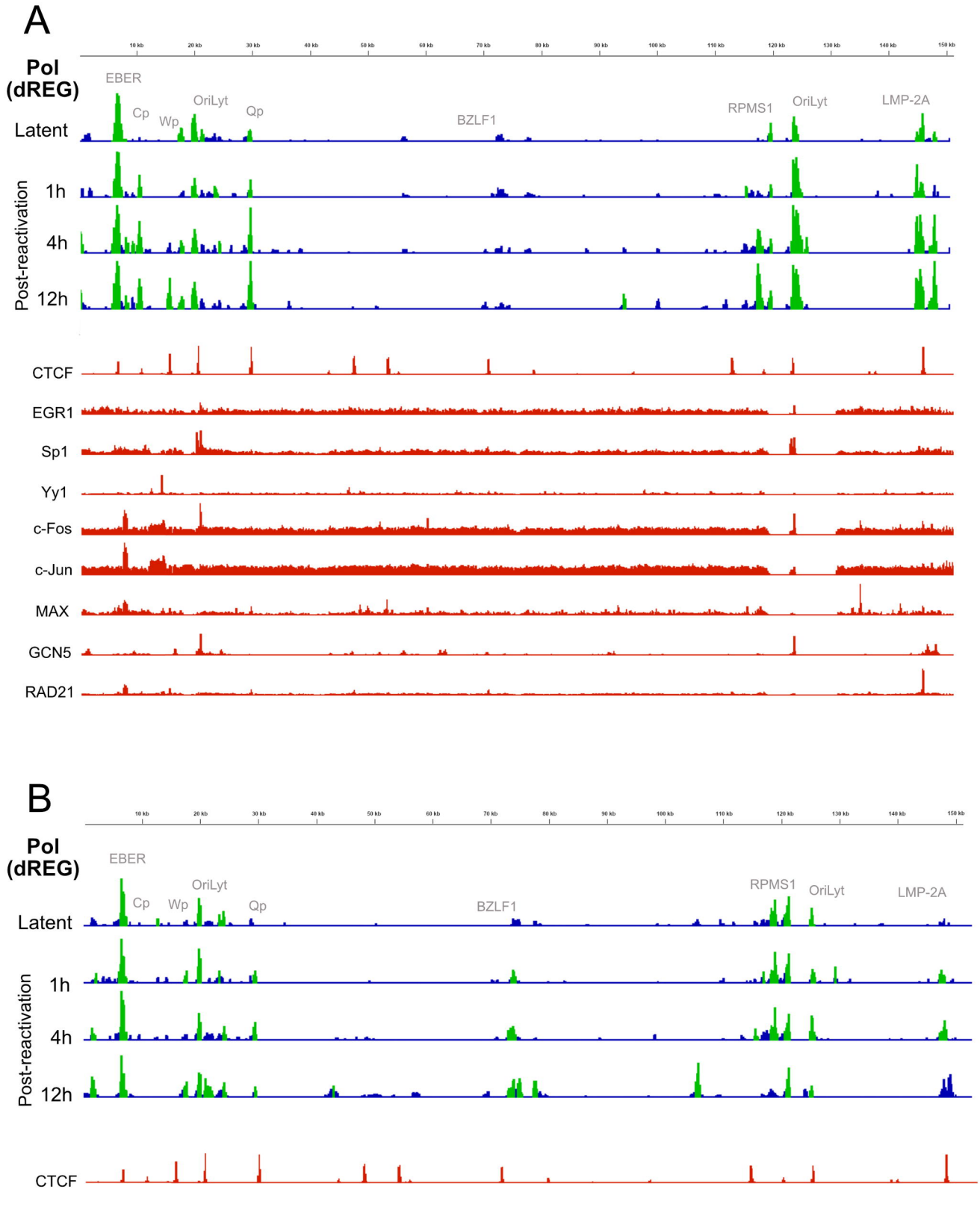

A
Latent
1h
4h
12h
Post-reactivation
Normalised Read Count
EBER
Cp
Wp
OriLyt
Qp
BHLF1
BRLF1
BZLF1
RPMS1
LMP-2A

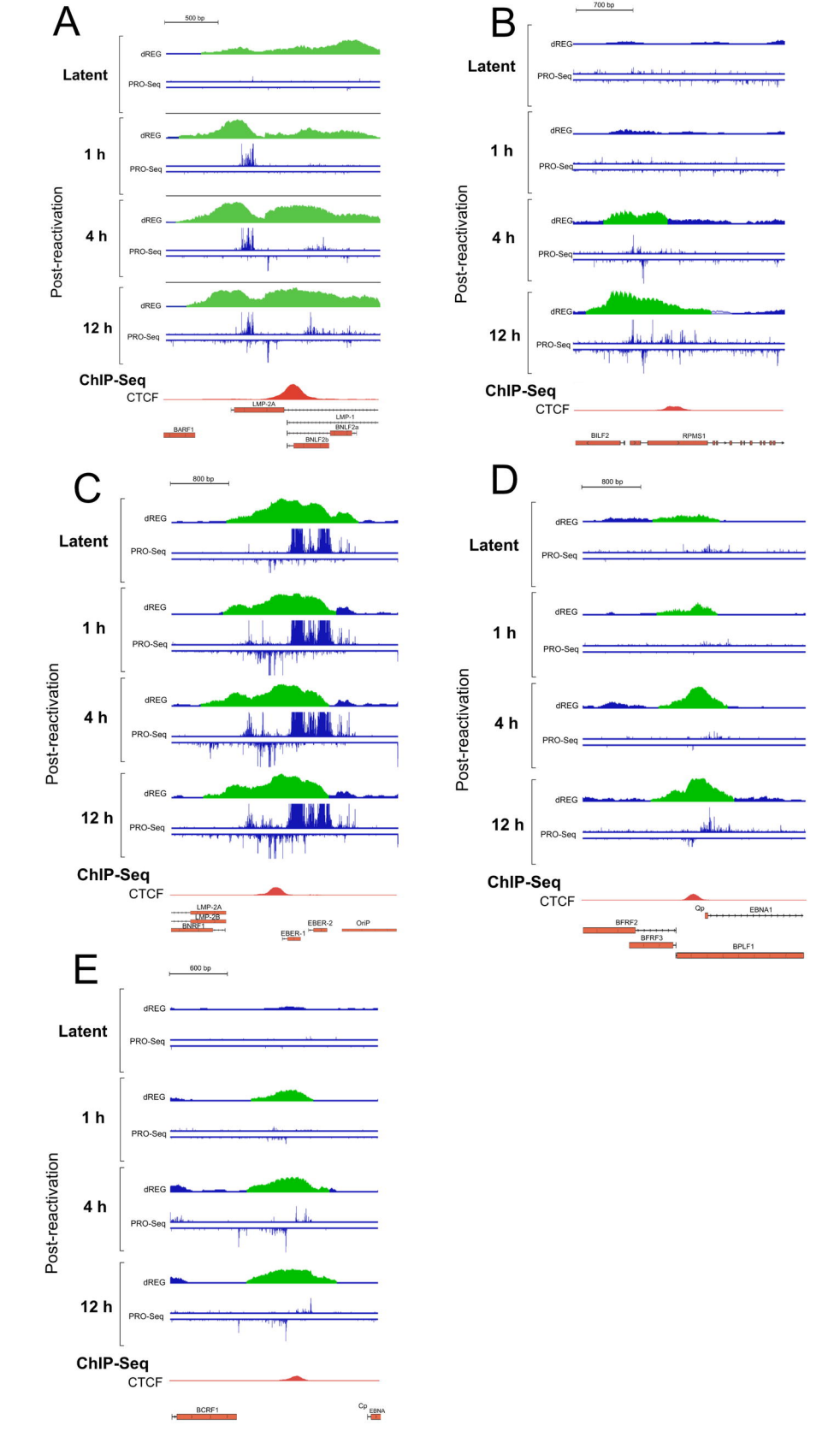

B
Latent v 1h post-reactivation
LMP-2A
EBER-1
-Log10(adjusted p-value)
Log2 Fold Change

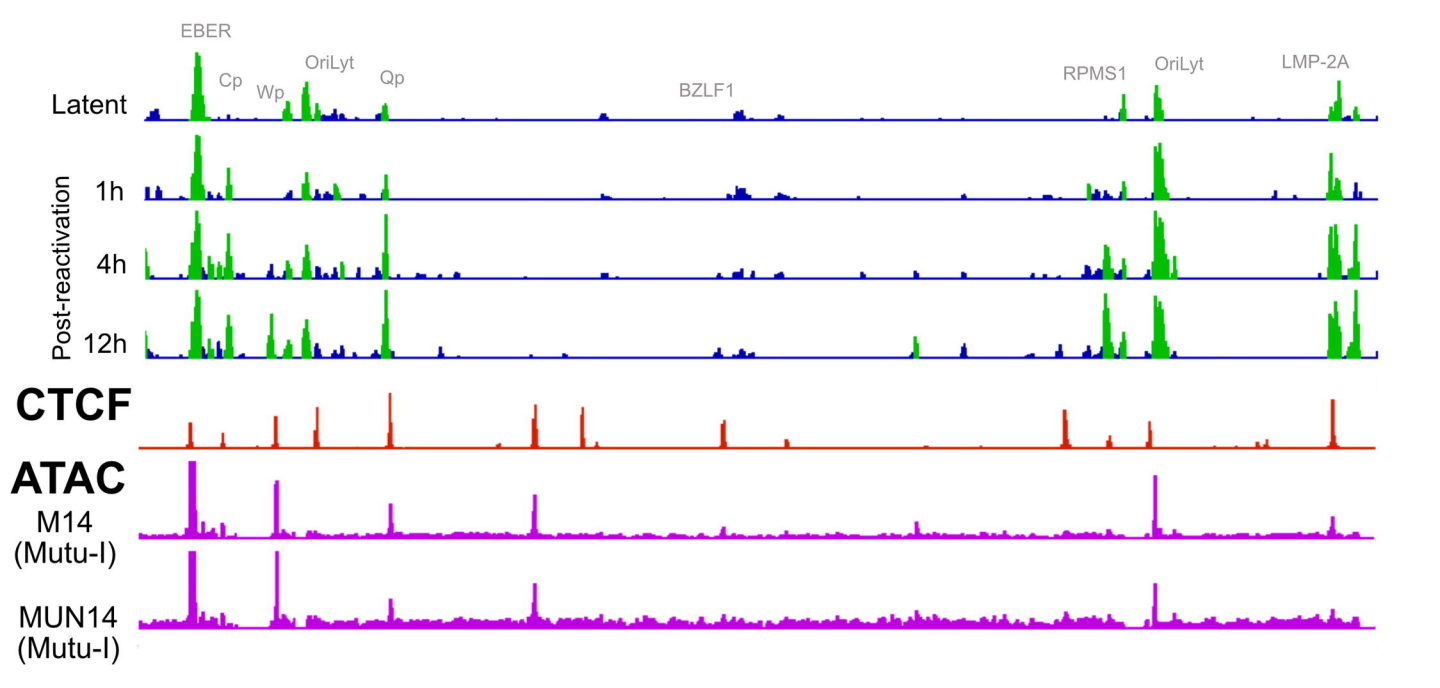

C
Latent v 4h post-reactivation
LMP-2A
RPMS1
LMP-2B
LMP-1
-Log10(adjusted p-value)
Log2 Fold Change

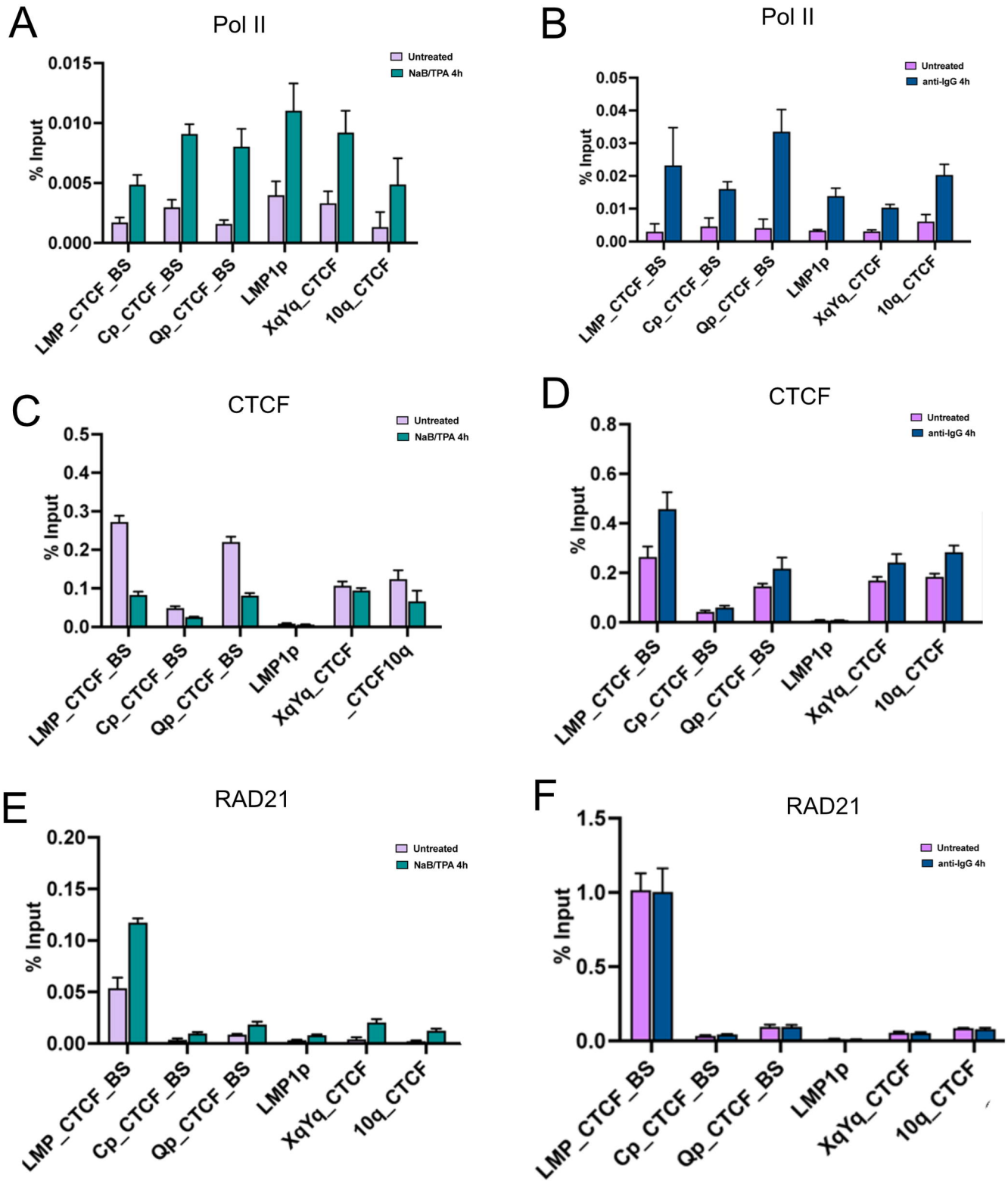

D
Latent v 12h post-reactivation
RPMS1
LMP-2A
BRLF1
BZLF1
-Log10(adjusted p-value)
Log2 Fold Change


A
1801 bp
Latent
Post-reactivation
1h
4h
12h
Normalised Read Count
50
0
LMP-2A
LMP-1
BNLF2a
BNLF2b
BARF1
B
1980 bp
Latent
Post-reactivation
1h
4h
12h
Normalised Read Count
EBER-2
OriP
EBER-1
C
11,000 bp
Latent
Post-reactivation
1h
4h
12h
Normalised Read Count
BVRF2
BILF2
RPMS1
BdRF1
LF3
OriLyt


A
10 kb
20 kb
30 kb
40 kb
50 kb
60 kb
70 kb
80 kb
90 kb
100 kb
110 kb
120 kb
130 kb
140 kb
150 kb
Pol (dREG)
EBER
Cp
Wp
OriLyt
Qp
BZLF1
RPMS1
OriLyt
LMP-2A
Latent
Post-reactivation
1h
4h
12h
CTCF
EGR1
Sp1
Yy1
c-Fos
c-Jun
MAX
GCN5
RAD21
B
10 kb
20 kb
30 kb
40 kb
50 kb
60 kb
70 kb
80 kb
90 kb
100 kb
110 kb
120 kb
130 kb
140 kb
150 kb
Pol (dREG)
EBER
Cp
Wp
OriLyt
Qp
BZLF1
RPMS1
OriLyt
LMP-2A
Latent
Post-reactivation
1h
4h
12h
CTCF


A
500 bp
Latent
dREG
PRO-Seq
1 h
4 h
12 h
Post-reactivation
ChIP-Seq
CTCF
LMP-2A
LMP-1
BARF1
BNLF2a
BNLF2b
B
700 bp
BILF2
RPMS1
C
800 bp
LMP-2B
BNRF1
EBER-1
EBER-2
OriP
D
Cp
EBNA1
BFRF2
BFRF3
BPLF1
E
600 bp
BCRF1
EBNA


EBER
Cp
Wp
OriLyt
Qp
BZLF1
RPMS1
OriLyt
LMP-2A
Latent
Post-reactivation
1h
4h
12h
CTCF
ATAC
M14
(Mutu-I)
MUN14
(Mutu-I)


A
Pol II
Untreated
NaB/TPA 4h
% Input
0.015
0.010
0.005
0.000
LMP_CTCF_BS
Cp_CTCF_BS
Qp_CTCF_BS
LMP1p
XqYq_CTCF
10q_CTCF
B
Pol II
Untreated
anti-IgG 4h
% Input
0.05
0.04
0.03
0.02
0.01
0.00
LMP_CTCF_BS
Cp_CTCF_BS
Qp_CTCF_BS
LMP1p
XqYq_CTCF
10q_CTCF
C
CTCF
Untreated
NaB/TPA 4h
% Input
0.5
0.4
0.3
0.2
0.1
0.0
LMP_CTCF_BS
Cp_CTCF_BS
Qp_CTCF_BS
LMP1p
XqYq_CTCF
_CTCF10q
D
CTCF
Untreated
anti-IgG 4h
% Input
0.8
0.6
0.4
0.2
0.0
LMP_CTCF_BS
Cp_CTCF_BS
Qp_CTCF_BS
LMP1p
XqYq_CTCF
10q_CTCF
E
RAD21
Untreated
NaB/TPA 4h
% Input
0.20
0.15
0.10
0.05
0.00
LMP_CTCF_BS
Cp_CTCF_BS
Qp_CTCF_BS
LMP1p
XqYq_CTCF
10q_CTCF
F
RAD21
Untreated
anti-IgG 4h
% Input
1.5
1.0
0.5
0.0
LMP_CTCF_BS
Cp_CTCF_BS
Qp_CTCF_BS
LMP1p
XqYq_CTCF
10q_CTCF