RNA Binding Proteins Shape Latent KSHV Genomic Structure: Identification of KSHV Episome Tethering Sites on Host Chromosomes

Ashish Kumar
UC Davis School of Medicine
Yuanzhi Lyu
UC Davis School of Medicine
Yuichi Yanagihashi
Lifematics
Chanikarn Chantarasrivong
Lifematics
Vladimir Majerciak
National Cancer Institute https://orcid.org/0000-0003-2907-4583
Michelle Salemi
University of California, Davis
Kang-Hsin Wang
UC Davis
Frank Chuang
UC Davis
Ryan Davis
UC Davis https://orcid.org/0000-0001-7439-6683
Clifford Tepper
UC Davis https://orcid.org/0000-0001-7105-1102
Kazushi Nakano
Lifematics
Chie Izumiya
UC Davis
Michiko Shimoda
UC Davis
Ken-ichi Nakajima
UC Davis
Alexander Merleev
UC Davis
Zhi-Ming Zheng

Tumor Virus RNA Biology Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA

Mel Campbell

UC Davis https://orcid.org/0000-0002-8475-6686

Yoshihiro Izumiya (YIZUMIYA@UCDAVIS.EDU)

UC Davis School of Medicine https://orcid.org/0000-0002-9184-2603

Article

Keywords: RNA Binding Proteins, KSHV, Episome Tethering Sites

Posted Date: September 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-136834/v1

RNA Binding Proteins Shape Latent KSHV Genomic Structure: Identification of KSHV Episome Tethering Sites on Host Chromosomes

Ashish Kumar[1, #], Yuanzhi Lyu[1, #], Yuichi Yanagihashi[2, #], Chanikarn Chantarasrivong[2], Vladimir Majerciak[3], Michelle Salemi[4], Kang-Hsin Wang[1], Frank Chuang[5], Ryan R. Davis[6], Clifford G. Tepper[5, 7], Kazushi Nakano[2], Chie Izumiya[1], Michiko Shimoda[1, 7], Ken-ichi Nakajima[1], Alexander Merleev[1], Zhi-Ming Zheng[3], Mel Campbell[1,*], & Yoshihiro Izumiya[1,5,7,*]

[1] Department of Dermatology School of Medicine, University of California Davis (UC Davis), Sacramento, California USA
[2] Lifescience Division, Lifematics, Osaka, 541-0046 Japan
[3] HIV Dynamic and Replication Program, National Cancer Institute, NIH, Frederick, Maryland, USA
[4] Genome Center, Proteomics Core, Genome and Biomedical Sciences Facility, UC Davis, Davis, California, USA
[5] Department of Biochemistry and Molecular Medicine, School of Medicine, UC Davis, Sacramento, California USA
[6] Department of Pathology and Laboratory Medicine, School of Medicine, UC Davis, Sacramento, California USA
[7] Viral Oncology and Pathogen-Associated Malignancies Initiative, UC Davis Comprehensive Cancer Center, Sacramento, California USA

[#] These authors contributed equally.

* Correspondence: Mel Campbell, PhD
Address: UCDMC Research III, Room 3300D, 4645 2nd Avenue, Sacramento CA 95817
E. mail: mcampbell@ucdavis.edu
Phone: 916-734-7811

*Correspondence: Yoshihiro Izumiya DVM, PhD
Address: UCDMC Research III, Room 2200B, 4645 2nd Avenue, Sacramento CA 95817
E. mail: yizumiya@ucdavis.edu
Phone: 916-734-7253

Word counts:

Abstract: 183

Main text: 4,331

Abstract

In previous studies, we have shown that expression of a viral lncRNA, polyadenylated nuclear RNA (PAN RNA) is essential for inducible viral genomic looping and distal gene activation during Kaposi's sarcoma-associated herpesvirus (KSHV) reactivation. Here we identify viral lncRNA binding proteins, and show that an underlying molecular mechanism regulating the KSHV latency-lytic replication switch is via a viral lncRNA-CHD4 (chromodomain helicase DNA binding protein 4) interaction. Proximity biotin labeling, single cell transcriptomics, and siRNA screening along with complementation studies identified that CHD4's enzymatic activity silences viral gene expression by preventing transcription factory formation. Furthermore, Capture Hi-C, Cleavage Under Targets and Release Using Nuclease (CUT&RUN), and proteomics approaches together identify KSHV episome docking sites on host chromosomes and colocalization with a CHD4 protein complex, ChAHP, at epigenetically active genomic regions. PAN RNA binds and competes with CHD4 DNA binding *in vitro,* and KSHV episomes detached from these host genomic loci sites when reactivation is triggered. Our studies suggest that CHD4 exhibits strong repressor function by preventing inducible genomic looping, and is therefore important for the ability of KSHV to establish and maintain latency in a "poised" state at selected host genomic loci.

Introduction

Kaposi's sarcoma-associated herpesvirus (KSHV), discovered in 1994, is one of eight human herpesviruses (1). Since then, KSHV has been identified as the causative agent of Kaposi's sarcoma (1-3) and two human lymphoproliferative diseases, primary effusion lymphoma (PEL) and AIDS-related multicentric Castleman's disease (4-7). Although HAART (highly active antiretroviral therapy) of HIV-patients dramatically reduced KS incidence in Western countries, KSHV associated-malignancies are still a leading cause of cancer deaths in AIDS patients in Sub-Saharan Africa. In cancer cells, KSHV establishes a latent infection in which most of the viral genes are silenced. External stimuli trigger KSHV reactivation from latently infected cells (8, 9).

Spatial and temporal organization of the genome play critical roles in gene regulation. Temporal bending of a promoter to contact active enhancers is an essential step for gene expression (10). Through chromosome conformation capture analyses with deep sequencing (Capture Hi-C; CHi-C), we found that the structure of the KSHV genome shifts to increase genomic looping at K-Rta binding sites during KSHV reactivation (11). The viral long non-coding RNA (Poly Adenylated Nuclear RNA; PAN RNA) is a major K-Rta transcriptional target, and this genomic locus recruits 40-fold more cellular RNA polymerase II (RNA Pol II) molecules than the cellular *ACTB* gene, which is consistent with an exponentially higher copy number of PAN RNA molecules in reactivating cells (12). Imaging studies also showed that KSHV genomes assembled to form "transcriptional factories", in which a significant fraction of RNA Pol II was recruited to KSHV genomes during reactivation (12).

The high copy number of PAN RNA has been partly explained by its higher RNA stability (13). At least three sequence elements, ENE (expression and nuclear retention element) at the

3' region, MRE (Mta responsive element) at the 5' region of PAN RNA, and the structure of the poly (A) tail, are critical for PAN RNA stability (14-16). PAN RNA also interacts with multiple cellular and viral proteins (15, 17) , and functions as a transactivator by sequestering repressors and/or recruiting histone modifying enzymes (18-20). Expression of PAN RNA is also involved in increasing the efficacy of protein translation by interacting with poly(A)-binding protein C1 (21). A recent study demonstrated that PAN RNA can be compensated by other viral lncRNA sequences without a significant loss of KSHV replication, which suggests the presence of a sequence-independent function.

The KSHV ORF57 protein, also known as mRNA transcript accumulation factor (Mta), is an immediate-early gene product, which interacts with a number of RNA binding proteins and RNAs (22). ORF57 protein functions affect most aspects regulating RNA-mediated biology, including increasing splicing efficacy and RNA stability, and facilitating mRNA translation (22), and inhibiting the formation of cellular RNA granules (23, 24).

lncRNAs associate with specific sets of RNA-binding proteins (RBPs) to form functional ribonucleoprotein (RNP) complexes. Some of the lncRNAs localize to particular nuclear bodies; for example, XIST localizes to the Barr bodies (inactive X chromosome), NEAT1 in paraspeckles, MALAT1 in nuclear speckles, TUG1 in polycomb bodies, and SATIII in nuclear stress bodies (25, 26). Notably, RNA molecules are known to induce the formation of phase separation through intrinsically disordered regions (IDRs) of neighboring proteins in a context-dependent manner (27-29). An idea for the functional mode of lncRNAs in forming a non-membranous ‘ribonucleoprotein (RNP) milieu’ through recruitment of protein IDRs has recently emerged; this mechanism provides a flexible, inducible, and dynamic molecular platform in 3D nuclear space for miscellaneous components to assemble at sites of transcription. Identification

of proteins involved in the regulation of the complex formation is therefore an important step towards understanding the molecular mechanisms of aggregation and their biological functions. Recent studies identified an RNA-binding domain in the CTCF protein, which regulates formation of topologically associating domains (30, 31), suggesting the presence of distinct classes of RNA-binding protein (RBP)-dependent genomic loops, which are likely regulated by nascent RNAs and other RBP partners.

In this study, we combined proximity biotin labeling with a recombinant bacterial artificial chromosome (BAC)-based KSHV manipulation system to probe for PAN RNA-binding proteins during reactivation. Our studies identified chromodomain helicase DNA binding protein 4 (CHD4), a component of the ChAHP complex, as a strong KSHV silencer by preventing induction of genomic looping and RNA pol II aggregate formation during reactivation. Importantly, KSHV episomes anchored on host chromosomes with the ChAHP complex at selective sites. We propose a model in which KSHV maintains latent chromatin in a "poised" state by utilizing the ChAHP complex to prevent enhancer-promoter interactions at epigenetically active sites.

Results

KSHV RNA binding protein functions in static and inducible genomic looping. Herpesvirus family encodes ORF57 (e.g., ICP27) homologs. An ORF57 knock-out KSHV has been generated and studies showed that ORF57 is essential for KSHV reactivation/replication (32). The study also showed that supplementing ORF57 cDNA in *trans* did not completely rescue KSHV replication from ORF57-knock-out BAC transfected cells. The results suggested the idea that ORF57 protein, like cellular RNA-binding proteins (30, 31, 33), may play a role in organizing KSHV genomic structure during the initial stage of *de novo* infection. Lack of ORF57 protein

during initial burst of transcription may make viral genomic structure less responsive to stimuli. Accordingly, we performed CHi-C in order to examine static and dynamic genomic loop formation in *i*SLK cells stably-transfected with wild type (Wt) or ORF57-Stop recombinant KSHV BAC DNAs. ORF57 protein deletion along with the expression of other viral proteins was confirmed by immunoblotting (Fig. 1a). The CHi-C results showed that ORF57-Stop mutation decreased long-range genomic loops (Fig. 1b right panel). However, latent gene cluster regions (i.e., located around 124-130 kb region) showed very similar frequencies and directionality of topologically associated domains (TADs) with those of wild type (Fig.1b, bottom panels). PAN RNA genomic regions have the highest gene density in latent cells, indicating these regions were most frequently neighboring other KSHV genomic fragments through looping-mediated contacts (Fig 1c), while latent gene cluster regions seem to avoid contacts with other viral genomic regions (Fig. 1c). Consistent with previous studies that demonstrated ORF57 regulates late gene expression, the genomic contacts most significantly affected by the ORF57-Stop were the late gene cluster regions, where genomic contacts with the immediate-early or early gene clusters (regions with active histone marks) were lost in ORF57-Stop cells (Fig. 1c, blue arrows). Differences in induced genomic loops further confirmed that genomic interactions between late gene clusters and IE/E clusters at 24 h post-reactivation were reduced by ORF57-Stop mutation (Fig 1d bottom panel; intersections show darker blue). However, the relative abundance of genomic loops inside TADs was instead increased in ORF57-Stop viral genomes. Reduced genomic contacts led viral gene transcription decline more quickly (Fig. 1e), resulting in a lower KSHV virion production in the culture supernatant (Fig. 1f). These results indicate that the viral RNA binding protein, ORF57 protein, has a function in shaping and/or maintaining viral genomic structure to optimize robust viral gene expression.

ORF57 is required to form nuclear aggregates associated with PAN RNA expression. KSHV formed nuclear aggregates during reactivation, and we described the structures as transcription factories, because of the colocalization with RNA pol II (12). Transcription factory is an aggregate of RNA, DNA, and protein molecules, and it is flexible structure for gene transcription (34). As ORF57 protein is known to interact with PAN RNA, we asked if ORF57 protein plays a role in the formation of nuclear aggregates with PAN RNA. The sizes and numbers of the nuclear aggregates were quantified after visualizing PAN RNA with RNA-FISH, and cellular and viral proteins with immunofluorescence staining. The results showed that cells harboring ORF57-Stop KSHV exhibited few RNA pol II aggregates while those with ORF57-Wt showed that a significant fraction (>98.6%) of PAN RNA positive cells contained many RNA pol II aggregates. K-Rta protein, but not ORF57 protein, colocalized with RNA pol II in PAN RNA expressing cells, dismissing requirement for ORF57 physical presence at transcriptional factories (Fig. 2a). There was a positive correlation between K-Rta signal intensity and PAN RNA expression in ORF57-Wt cells (Extended Data Fig. 1). However, such correlation was diminished in the absence of the ORF57, and overall PAN RNA signals were significantly reduced in ORF57-Stop cells (Extended Data Fig. 1). The Cleavage Under Targets and Release Using Nuclease (CUT&RUN) studies demonstrated marked enrichment of RNA Pol II on the KSHV genome during reactivation in ORF57-Wt KSHV (blue line) but not in ORF57-KO cells (Fig. 2b bottom panel, blue vs. pink line). The results were in good agreement with the aggregation of RNA pol II in imaging studies (Fig. 2a), and suggest that maintaining and/or recruitment of RNA Pol II on the KSHV genome required ORF57 protein. Studies with lncRNA (PAN RNA)-KO (11), which phenotype-copied ORF57-Stop, also suggest that expression of robust viral lncRNA is important for transcription factory formation. To further confirm the

significance of PAN RNA expression in distal promoter activation, we performed a reporter assay, using a bicistronic construct that encodes a PAN RNA expression cassette driven by K-Rta with a downstream ORF16 promoter controlling luciferase expression. A series of mutations were also inserted in the reporter to examine the molecular action of the viral lncRNA (Fig. 2c). The results showed that in the presence of wild type PAN RNA, the ORF16 promoter was synergistically activated by K-Rta and ORF57 protein. This synergistic activation was lost by mutation of K-Rta responsive elements in the PAN RNA promoter or by insertion of a poly(A) signal immediately after the PAN RNA transcription initiation site (Fig. 2c). Deletion of the nuclear retention element (ENE) (14) did not affect ORF16 promoter activation significantly, indicating that the action of transcription in *cis* may play a larger role in distal promoter activation. These synergistic effects required PAN RNA expression, since deletion of the PAN RNA cassette eliminated the ORF16 promoter activation completely, and the ORF57 nuclear localization signal mutant also failed to enhance ORF16 promoter activation (Fig. 2c right panels). Together, these results confirmed that PAN RNA expression is important for activation, or de-repression of distal promoter(s).

Identification of lncRNA-binding dependent protein-protein interaction by proximity biotin labeling. To understand the roles of ORF57 protein in distal promoter activation, we applied the proximity labeling technique to profile PAN RNA neighboring protein during KSHV reactivation (35). The 3xFlag-mini-TurboID (mTID) cassette was inserted as an N-terminal fusion of ORF57, a PAN RNA binding protein, thus the fusion protein is expressed from the endogenous promoter during KSHV reactivation (Fig. 3a). Taking advantage of previous detail mapping study, which identified ORF57 binding sites on PAN RNA, termed MRE (Mta Responsive

Element)(15), we utilized a PAN MRE mutant virus having a 9-nucleotide mutation in the MRE element (Fig. 3a). KSHV genome-wide qRT-PCR array analysis showed that the PAN RNA MRE mutation impaired viral gene expression globally, with stronger effects being seen again in late gene cluster regions (Fig. 3b, right panel). Protein biotinylation happened only in the presence of both ORF57 protein (reactivated sample) and biotin, which confirmed specificity (Fig. 3c). By comparing the MRE mutant with Wt, we narrowed cellular proteins responsible for increasing PAN RNA transcripts. Proximity biotin labeling identified a total of 129 and 307 proteins from mTID-57 PAN Wt *i*SLK cells and mTID-57 PAN-MRE *i*SLK cells, respectively, as specific ORF57-interacting proteins with p values lower than 0.05 (Extended Data Table 1). Among the interacting proteins, 74 proteins were in common between the PAN MRE Wt and PAN MRE mutant, while 55 proteins were found only in the presence of wild type PAN RNA sequence (Fig. 3d, e). Deletion of MRE seemed to unleash ORF57 protein and allows ORF57 to interact more freely with other RNA binding proteins (Fig. 3d). Gene Ontology (GO) analysis indeed showed that the ORF57-interacting proteins are primarily involved in mRNA processing (Fig. 3f).

Chromodomain helicase DNA binding protein 4 (CHD4) is a strong KSHV gene silencer.

Having a list of cellular proteins that may play a role in the establishment of nuclear aggregates, we next performed siRNA-based screening of ORF57-neighboring proteins for functions in gene transcription using KSHV lytic gene expression as readout (Fig. 4a). Among the 55 cellular proteins that were specifically associated with PAN RNA containing a functional MRE, 12 cellular proteins suppressed KSHV gene expression more than 2-fold. Especially, knock-down of the most highly-enriched protein, CHD4, enhanced KSHV reactivation more than 5-fold, and CHD4 was the strongest repressor among 129 proteins (Fig. 4a). Conversely, functional re-introduction

of CHD4 by over expression of mouse *Chd4* cDNA (i.e., in order to escape from the shRNA, which targets human CHD4) counteracted effects of CHD4 knockdown (Fig. 4b). Notably, over expression of mouse Chd4 almost completely abolished K-Rta mediated KSHV reactivation (Fig. 4b), and inhibited the transcription factory formation (Fig. 4c). Strong silencing effects were helicase-activity dependent, since mutations in the helicase domain, which were identified in CHD4-associated syndrome (36, 37), failed to inhibit KSHV reactivation (Fig. 4c, Extended Data Fig. 2). Restriction of KSHV reactivation by CHD4 could also be seen at the single cell level. Single cell RNA-sequencing of 24 hrs post-stimulated sample demonstrated higher levels of KSHV transcripts in CHD4 low cells (Extended Data Fig. 3).

PAN RNA inhibits CHD4 DNA binding. PAN RNA MRE-dependent biotinylation by ORF57 suggested that CHD4 is an RNA-binding protein and that PAN RNA may bring the two proteins into proximity by serving as a scaffold (model shown in Fig. 3a). We have purified these three components (Fig. 4d) and tested the possibility in isolated reactions. The results showed that CHD4 was indeed precipitated with RNAs in the condition, which both DNA binding protein, NF-κB (p65) and Luciferase were not precipitated with RNAs. The CHD4 RNA-binding seemed sequence-independent and also did not require ORF57 protein (Fig. 4e). CHD4 was recruited to the KSHV genome mostly at TAD boundaries with active histone marks (H3K4me3), which includes three lncRNA coding regions (Fig. 4f). In addition, the CHD4 was clearly enriched at LANA binding sites such as terminal repeat region in naturally infected PEL cells (Fig. 4f, Extended Data Fig. 4). Consistent with a previous study that Drosophila melanogaster CHD4 homolog is capable of binding DNA and human CHD4 ATPase activity is stimulated in presence of naked DNA (38), purified CHD4 protein bound to a double-stranded (ds) DNA fragment

encoding the PAN RNA sequence (Fig. 4g). Importantly, presence of increased amounts of PAN RNA (ssRNA) clearly antagonized the CHD4 dsDNA binding, which was completely blocked at 1:10 (dsDNA/RNA) molecular ratio (Fig. 4g).

KSHV episomes tether host chromosomes with ChAHP. Very strong silencing effects of CHD4 on KSHV lytic gene expression and TR localization (Fig. 4b, Extended Data Fig. 4) prompted us to examine the association of CHD4 binding sites on host cellular chromatin with KSHV episome tethering sites. Capture Hi-C (CHi-C) analysis was employed to identify the genomic loci which exhibited higher frequencies of chimeric ligation reads formed between KSHV DNAs and host genomic fragments. The results demonstrated that KSHV genomic CHi-C sequence reads were highly enriched at selected host chromatin loci (Fig. 5a, individual chromosomes were presented in Extended Data Figure 5a-c). Three naturally-infected PEL cell lines demonstrated a very similar distribution of tethering sites, yet there were also cell line-specific sites (Fig. 5a, b). For example, BC-1 and BC-3 had major episome binding sites on chromosome 12, but BCBL-1 lacked those peaks (Fig. 5a). Zooming in on individual chromosomes further showed that KSHV sequence reads were frequently seen near the centromeric region (Fig. 5b right panel, Extended Data Figure 5a-c). Overlaying LANA CUT&RUN sequence reads further demonstrated that CHi-C enriched sites overlapped with LANA CUT&RUN signals (Fig. 5c, Extended Data Fig. 6); this result is consistent with LANA serving as an anchor protein for chromosome tethering (39-41). We noted that there were two types of LANA CUT&RUN signals; one with sharp peaks characteristic of cellular transcriptional factor binding, and the other showing much stronger/higher peaks that often spanned more than 10 kb (Fig. 5c, left panel). Because LANA is known to bind 28 kb of KSHV terminal repeat

sequences (i.e., 0.8 kb/repeat x 35 copies of TR), and those larger peaks clearly overlapped with CHi-C signals, we considered that those host genomic regions are the KSHV episiome tethering sites. To our surprise, these KSHV tethering areas were also major CHD4 recruitment sites (Fig. 5c, left bottom panel). We separated the peaks into two groups based on the peak distance spanned (e.g., >1 kb or default setting for MACS2 analysis) and then performed separate statistical analyses. The pairwise intersection-Jaccard statistic showed that the TR bound form of LANA was clearly colocalized with CHD4 in experimentally-infected *i*SLK cells as well as BCBL-1 with Jaccard similarity coefficients of 0.69 and 0.42, respectively (Fig. 5d, Extended Data Figure 7). Venn diagram showed that CHD4 had 28 peaks that spanned more than 1 kb, and 27 of these CHD4 expanded-peaks colocalized with LANA peaks in BCBL-1 cells (Fig. 5d). As expected, acetylated histone (H3K27Ac) and RNA pol II had a larger number of such extended peaks along host chromosomes, and LANA binding sites were often colocalized with RNA pol II and active histone marks (e.g., 40 out of 64 sites; Fig. 4f, Fig. 5d). Co-occupancy analyses also showed that CHi-C reads were primarily localized at active genomic regions but not with H3K36me3 or H3K27me3 marks (Extended Data Figure 8). Interestingly, comparison of CHi-C contacts with reactivated samples indicated that reactivation detached KSHV episomes from host chromosomes (i.e., reducing relative contact reads with host chromatin; Fig. 5e). DNA-FISH performed with centromere-specific PNA (peptide nucleic acid) probes in combination with LANA immunostaining confirmed that the multiple LANA dots were closely localized with centromeres in the nucleus, further supporting the CHi-C results (Fig. 5f, Extended Data Figure 9). In addition, ENCODE atlas data obtained from high resolution Hi-C projects conducted with other cell lines indicated that these KSHV episome tethering sites have higher frequencies of genomic loops, indicating the genomic regions possesses characteristics of gene enhancers

(Extended Data Fig. 6). Taken together, the results suggested that KSHV episomes are tethered to unique genomic loci (e.g. active genomic regions near heterochromatin), where locally concentrated CHD4 may function to silence active incoming episomes for latency establishment.

LANA interacts with CHD4. We next performed proximity biotin labeling with recombinant mini-mTID-LANA virus to examine local chromatin environment of LANA. With this, we confirmed that KSHV LANA protein indeed localizes in close proximity to CHD4 in latently infected cells, supporting CUT&RUN results (Fig. 5g, Extended Data Table 2). Importantly, the proteomics studies did not identify NuRD complex components, such as MTAs or HDACs. Instead, the study identified components of the ChAHP complex, another CHD4 protein complex, ADNP (Activity-dependent neuroprotector homeobox protein) and HP-1γ interacting protein with high confidence (Fig. 5g), suggesting that episome-tethering regions are likely to be regulated by ChAHP but not NuRD complex. Pull-down studies with purified proteins demonstrated that LANA could interact with CHD4 in the absence of other molecules (Fig. 5h).

CHD4 is important for KSHV latency establishment. The studies above suggest that CHD4 silences KSHV transcription, which consequently supports the establishment of KSHV latency. We tested this hypothesis by knocking-down CHD4 in 293FT cells (Fig. 6a) and then infected these cells with KSHV to monitor viral gene silencing by collecting total RNA. Purified KSHVr2.19 was used for *de novo* infection with a MOI of 1. The results showed that KSHV gene expression continued to increase during a 3-day period in two independent CHD4 KD cells, but not in control shScramble cells (Fig. 6b). The results suggested that CHD4 plays an essential role in KSHV

latency establishment (Fig. 6b) and maintenance (Fig. 4b) by silencing KSHV lytic gene expression.

Discussion

Recent studies indicated an RNA-binding domain in transcription factor CTCF is essential for the formation of TAD subsets (30, 31). The presence of distinct classes of RNA-binding protein (RBP)-dependent genomic loops suggests the partition of nascent RNAs and other RBP partners in chromatin-looping at transcriptionally active gnomic loci. KSHV ORF57 protein is expressed during KSHV lytic replication, and we found that ORF57 expression affects genomic looping of the latent episomes via expression of viral lncRNA. The results suggested that the 3D latent genomic structure is regulated by lytic transcription program mediated by immediate-early proteins through the recruitment of factors or nascent RNAs.

It is known that enhancer RNAs are part of transcriptional factory formation (42-44), and we propose that CHD4 enzymatic activity contributes to preventing the DNA/RNA aggregates and therefore inhibits recruitment of RNA-binding proteins (30); this action prevents RNA pol II for accessing other genomic regions and inhibits "recycling" of assembled RNA pol II complex (Fig. 2c). Notably, loss of function mutations in CHD4 are associated with multiple developmental disorders (45-50). We think that such losses may result in increased sizes of transcription factories and propensity to form active enhancer-promoter interactions, which would prolong gene transcription; it was clearly seen in KSHV-infected CHD4-KD 293FT cells (Fig. 6b).

Our studies agree with a recent report, which demonstrated that conditional knock-out of CHD4 in brain tissue increases enhancer accessibility and cohesin binding to promote gene

expression (51). The conditional CHD4 knock-out strengthens domain contacts and looping between promoters and enhancers (51). CHD4 regulation of super-enhancer accessibility in rhabdomyosarcoma has been found in association with cancer vulnerability (52). Importantly, in an epigenetic factor screening, CHD4 was identified as a strong repressor to KSHV replication (53). However, how KSHV episomes select specific CHD4 binding sites from many CHD4 bound genomic regions or whether the selected CHD4 sites are linked to cell differentiation, tissue-specific gene expression and therefore KSHV tissue tropism remain unknown. It would be also interesting to know if KSHV episome-bound “enhancers” become under the control by KSHV proteins in infected cells.

KSHV evolved to maintain a mysterious lncRNA, which is expressed at very high copy number in the presence of ORF57 protein, and deletion of PAN RNA or ORF57 significantly impairs the entire viral gene expression program (11, 18, 32). How does PAN RNA expression activate the expression of other viral open reading frames? Here, we provided evidence that PAN RNA binds CHD4 and relieves CHD4 from dsDNA binding as a transcription suppressor protein, thereby promoting viral lytic gene expression. Two viral immediate-early proteins RTA and ORF57 synergize the effect by RTA activating the PAN promoter for transcription and ORF57 stabilizing the transcribed PAN RNA. We speculate that very high copies of PAN RNA may overwhelm and sequester away CHD4 from the KSHV genomic enhancer regions. Because PAN RNA expresses over 10^5 copies, and CHD4 localizes near PAN RNA transcription initiation sites with LANA (Fig. 4f), we expect local PAN RNA concentration would easily result in an RNA-DNA ratio in excess of 10:1 (Fig. 4g). Based on frequent genomic looping, non-coding RNA expression, and CHD4 binding (e.g., which frequently targets enhancers), we also propose that the lncRNA encoding regions [Ori-RNA (T1.5), PAN RNA, K12 (T0.7)] act as inducible

enhancers for the KSHV ORFs, and KSHV reactivation (e.g., promoter activation) hinges on lncRNA-CHD4 interactions for de-repression. Similar to CHD4, we previously demonstrated that PAN RNA expression also sequesters LANA from unique regions of KSHV genomes (19). Our CUT&RUN studies showed CHD4 and LANA colocalized on both the KSHV and host genomes (Fig. 4f) and they could physically interact each other (Fig. 5h); thus, we favor the idea that reactivation is triggered by detachment of LANA/CHD4 complex-loaded TR fragments from the KSHV unique region (ORFs-encoded region) by robust lncRNA expression. Gene-density analysis indeed indicated such action, where TR unit rapidly lost neighboring-DNA fragments during reactivation (Fig. 2C left panel). Noteworthy, our KSHV 3D genomic structure modeling suggested that two highly inducible lncRNA encoding regions [PAN RNA, K12 (T0.7)] localizes close proximity to TR region (Campbell et al., submitted). The results combined with this study may suggest that CHD4-LANA complex has an architectural role for KSHV latent genomic structure, and TR region that recruits a significant amount of both CHD4 and LANA (Extended Data Figure 4), is important for structurally and epigenetically. Disruption of such a "backbone" by CHD4 KD would therefore unleash a viral lytic gene transcription program (Fig. 4b, 6b). Further studies will be needed to prove this.

The significance of CHD4 in maintaining KSHV latency is also supported at the single cell level. The probability of KSHV reactivation (KSHV transcripts) and CHD4 expression appears to be negatively correlated, but such negative correlation was not seen with similarly expressed another NuRD complex component, HDAC2. Proximity labeling with mini-TurboID LANA also did not identify HDACs or MTAs (Fig. 5g). Interestingly, Ostapcuk et al. reported another stable CHD4 containing protein complex, ChAHP, which consists of ADNP (activity-dependent neuroprotective protein), HP-1β/γ, and CHD4 (54). We identified ADNP, CHD4, and HP1 binding

protein 3 in proximity of LANA (Fig. 5g), suggesting that KSHV episomes may target ChAHP binding sites or TR regions attract ChAHP complex through multiple copies of CTCCC core sequences in a TR unit (54, 55). Furthermore, ChAHP is reported to compete with CTCF binding, and therefore counteracts chromatin looping at CTCF binding sites and maintain evolutionarily conserved spatial chromatin organization by preventing new CTCF bindings that emerged through short Interspersed element expansions (55); this suggests that KSHV cleverly find/build a “safe basecamp” in host chromosomes to maintain viral episomes during evolution (Extended Data Fig. 10). A recent study from Dr. Lieberman's group demonstrated that EBV episomes are tethered in the neighborhood of transcriptionally silent neuronal genes (56). Our C-HiC also found KSHV episomes frequently localized junctions of between eu- and hetero-chromatin (Extended Data Fig. 10 - junction of stripes).

In summary, we have discovered that the gene regulatory ChAHP complex is a key regulator of KSHV latency, and identified KSHV episome tethering sites on the host cellular chromosomes. Our studies suggested that CHD4 binding to DNA as a transcription suppressor to prevent genomic looping. However, PAN RNA expression mediated by RTA and ORF57 during lytic infection enhances viral gene expression by alteration of chromatin looping through its binding to CHD4 (Fig. 6c). With strong effects of CHD4 on KSHV replication, it will be important to study how ChAHP complexes are regulated and to determine the relationship with KSHV replication, including association of KSHV-mediated diseases with individuals who unfortunately possesses CHD4 mutations.

Methods

Materials. Dulbecco's modified minimal essential medium (DMEM), Fetal bovine serum (FBS), phosphate buffered saline (PBS), Trypsin-EDTA solution, 100x Penicillin-streptomycin-L-glutamine solution and Strepavidin-HRP conjugate were purchased from Thermo Fisher (Waltham, MA USA). Puromycin and G418 solution were obtained from InvivoGen (San Diego, CA, USA). Hygromycin B solution was purchased from Enzo Life Science (Farmingdale, NY, USA). Anti-ORF57, anti-K8α, and anti-K8.1, antibodies were purchased from Santa Cruz Biotechnology Inc (Santa Cruz, CA, USA). Anti-HA.11 epitope tag monoclonal antibody was obtained from Fisher scientific (Thermo Fisher). Anti-K-Rta antibody was described previously (57). All other chemicals were purchased from Millipore-Sigma (St. Louis, MO, USA) unless otherwise stated.

Cells, siRNA transfection and reagents. *i*SLK.219 cells were maintained in DMEM medium supplemented with 10% FBS, 1% penicillin-streptomycin-L-glutamine solution, 10 µg/ml puromycin, 400 µg/ml hygromycin B, and 250 µg/ml G418. *i*SLK cells were obtained from Dr. Don Ganem (Novartis Institute for Biomedical Research) and were maintained in DMEM supplemented with 10% FBS, 1% penicillin-streptomycin-L-glutamine solution and 10 µg/ml puromycin. BC-1 and BC-3 cell line was obtained from ATCC (Manassas, VA, USA), expanded and stored to obtain early passage stocks. BCBL-1 cell line was also obtained from Dr. Ganem (University of California San Francisco) in 2001. BC-1, BC-3, and BCBL-1 cell lines were cultured in RPMI 1640 medium supplemented with 15% FBS and 1% penicillin-streptomycin-L-glutamine solution. Transfection of siRNA in *i*SLK.219 cells was performed with Lipofectamine RNAiMax reagent (Thermo Fisher) according to manufacturer's protocol.

Capture Hi-C. KSHV Capture Hi-C (CHi-C) analysis was performed using a robust *in situ* CHi-C protocol with kitted reagents from Arima Genomics (San Diego, CA, USA) based on methods

described for *in situ* Hi-C (58, 59), CHi-C (60-62), and as described previously (11). Briefly, cells were crosslinked with 2% formaldehyde, lysed, and the genomic DNA digested with a cocktail of 4-cutter restriction endonucleases by incubation for 30 minutes at 37°C. The 5'-overhangs were then filled in and labeled with biotinylated dATP (biotin-14-dATP) by incorporation with Klenow fragment of DNA polymerase I (incubation for 45 minutes at 25°C). Ligation of the spatially proximal blunt-ended fragments was then performed with T4 DNA ligase (incubation for 15 minutes at 25°C). The formaldehyde crosslinks were reversed and the proximally-ligated, chimeric DNA products were purified with Agencourt AMPure XP paramagnetic beads (Beckman Coulter, Brea, CA, USA). The DNA was then fragmented to an average size of 400 bp with a Covaris E220 Focused-ultrasonicator (Covaris, Inc., Woburn, MA, USA) and size-selected to have a fragment size distribution of 200-600 bp with AMPure XP beads. The biotin-labeled ligation products were then selectively enriched by affinity capture with streptavidin magnetic beads (Arima Enrichment Beads). Subsequently, sequencing libraries were prepared with the Kapa HyperPrep Kit with Library Amplification Module (Roche, Basel, Switzerland) using a modified protocol for on-bead end repair, dA-tailing, and ligation of Illumina TruSeq sequencing adaptors.

The KSHV CHi-C library was then prepared from the Hi-C libraries as previously described (11). Briefly, target enrichment for KSHV genomic content was performed by solution hybridization with a custom-designed KSHV genomic capture probe library (xGen Lockdown Probes; Integrated DNA Technologies, Inc., Coralville, IA) (11) and subsequent capture of the hybridized targets with streptavidin beads (DynaBeads MyOne Streptavidin C1; Thermo Fisher) according to the manufacturer's standard protocol (Integrated DNA Technologies, Inc., Coralville, IA). The KSHV genome-enriched CHi-C library DNA was eluted and PCR enrichment

(12 cycles) performed with high-fidelity KAPA HiFi HotStart DNA Polymerase (Kapa Biosystems, Inc., Wilmington, MA). Libraries were multiplex sequenced (2 x 150bp, paired-end, ~50 million mapped reads/mate pairs per sample) on an Illumina Hiseq 4000 sequencing system.

Cleavage Under Targets and Release Using Nuclease (CUT&RUN). CUT&RUN (63) was performed by essentially following the online protocol established by Dr. Henikoff's lab with some modifications to fit our needs. Cells were washed with PBS and wash buffer (20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine (Sigma, S2626) and proteinase inhibitor (Roche). After removing wash buffer, cells were captured on magnetic ConA beads (Polysciences, PA, USA), in the presence of $CaCl_2$. Beads/cells complexes were washed with digitonin wash buffer (0.02% digitonin, 20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine and 1x proteinase inhibitor) 3 times, and incubate with specific antibodies in 250 μL volume. The antibodies used in this study were: mouse anti-RNA polymerase II (Millipore, clone CTD4H8; 1:100); rabbit monoclonal anti-CHD4 (Cell Signaling, D4B7; 1:50); rat monoclonal anti-LANA (Millipore-Sigma, clone LN53; 1:50). After incubation, unbound antibody was washed by digitonin wash buffer 3 times. Beads were then incubated with the recombinant pAG-MNase, which was purified from *E.coli* (Extended Data Fig. 11), in 250 μl digitonin wash buffer at 1.0 μg/mL final concentration for an hour at 4°C rotation. Unbound pAG-MNase was removed by washing with digitonin wash buffer 3 times. Pre-chilled 2 mM $CaCl_2$ containing digitonin wash buffer (200 μL) was added to beads and incubate on ice for 30 min. The pAG-MNase digestion was halted by the addition of 200 μl 2× STOP solution (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 50 μg/ml RNase A, 50 μg/ml glycogens). The beads were incubated with shaking at 37°C for 10 min in a tube shaker at 500 rpm to release digested DNA fragments from the insoluble nuclear chromatin. The supernatant was collected after centrifuge (16,000 x g for 5 min at 4°C)

and place on magnetic stand. DNA was extracted using NucleoSpin kit (Takara Bio, Kusatsu, Shiga, Japan). Sequencing libraries were then prepared from 3 ng with the Kapa HyperPrep Kit (Roche) according to the manufacturer's standard protocol. Libraries were multiplex sequenced (2 x 150bp, paired-end) on an Illumina HiSeq 4000 sequencing system to yield ~15 million mapped reads per sample. E. coli genomic DNAs from a pAG-MNase incubation was used to normalize data as described previously (63).

Genomic data analysis. The HiC-pro 2.11.1 pipeline (64) was used to align sequences from the Hi-C experiments against a combined assembly of reference genomes; the human hg19 (GRCh37) and KSHV (NC_009333.1). The reads were filtered to remove duplicates and to keep only reads that mapped to the KSHV genome by using the Python module Pysam 0.14.1. TADs, border strength, insulation scores, and density plots of the KSHV genome were analyzed using TADbit (65) following the developer's manual. The KSHV mapped reads were extracted for only uniquely mapped reads pairs by identifying intersection of each read-end, and the valid reads were provided by removing reads with self-circle, dangling-end, error, extra dangling-end, too short, too large, duplicated, and random breaks. The valid reads were stored as matrices and binned with resolution of 500 bp (2 kb for density plots); the bins with more than 100 counts and at least 75% of cells with no-zero counts were used in the next steps. Iterative Correction and Eigenvector decomposition (ICE) normalization was used to treat the data with iteration of 100. Insulation score/border strength was calculated using the TADbit algorithm (65) and visualized as a contact map. Gene density was plotted from the 3D model of each sample. Briefly, the Matrix Modeling Potential (MMP) score (66) was used to identify potential matrices, and the chosen matrices of all samples had scored more than 0.91. The parameters of the matrices, included maximal distance associated between two interacting particles, particles attraction,

particles repulsion, and contact distance of particles, were optimized using the Monte Carlo optimization method. Sets of models were produced from possible combination of those parameters, and used to build contact maps. The contact maps were then compared with the Hi-C interaction by averages of Spearman correlation coefficients. The models with higher correlation coefficients represented the original data. The model correlations of all samples in this experiment were higher than 0.79.

Fastq files for the Hi-C experiments were also processed through the HiCUP (v0.7.4) pipeline (67) using a combined human hg19 (GRCh37) and KSHV (NC_009333.1) genome. Valid interaction products called by HiCUP were converted into Juicebox (68) input format (.hic file), which stores the normalized and un-normalized contact matrices as a highly compressed binary file, by using a series a scripts provided by HiCUP (hicup2homer) and HOMER (makeTagDirectory and tagDir2hicFile) (69). Juicebox was utilized to facilitate adjustments of resolution and normalization, intensity scaling, zooming, and addition of annotation tracks.

CUT&RUN sequence reads were aligned to human hg38 reference genome and reference KSHV genome sequence (Human herpesvirus 8 strain: GQ994935.1) with Bowtie2 (70). MACS2 was used for detecting peaks (71) following the developer's manual. Peaks and reads alignments were visualized using the Integrated Genome Browser (IGB) (72). Furthermore, the peaks called by MACS2 with Q-value of 0.05 with sizes >1,000 bp were analyzed and visualized using Intervene v0.6.4 (73). Intersections of genomic peaks of 4 different protein binding sites were analyzed with default parameters and visualized as a Venn diagram. The similarity of positions of each protein pair binding sites was computed using the Jaccard statistic and illustrated as pie charts.

Single cell RNA sequencing data analysis. Single cell data was analysed with the Cellranger v2.1 pipeline (10x genomics). The pipeline included alignment to the hg38 human reference genome and human herpesvirus 8 strain (GQ994935.1) reference genome, t-distributed stochastic neighbor embedding (tSNE), and K-means clustering. Reads count matrix obtained from the pipeline were normalized using log normalization method with “Seurat” R package (74). To perform correlation analysis the KSHV genes expression was summarized, log transformed and resulted values divided by 14 equal intervals to have more than 50 cells in each interval. Median value of genes expression calculated for cells in each interval. The median values used to analyse and visualize correlation of cells and virus's gene expression.

Immunofluorescence staining analyses. *i*SLK cells latently infected with BAC16 ORF57-Wt were seeded onto glass coverslips and transfected with pLenti-Flag-mCHD4-myc WT or pLenti-Flag-mCHD4-myc mutant using Lipofectamine 2000 reagent according to the manufacturer’s protocol. pLenti-Flag-mCHD4-myc was obtained from commercial source (Origene, MD, USA). Point mutations were inserted by synthesizing DNA fragments that contain intended mutations and inserted at SexAI-AgeI restriction sites with In-FusionHD (Takara Bio, Japan). At 24h after transfection, 1 μg/ml of doxycycline and 20 ng/ml of 12-O-Tetradecanoylphorbol-13-acetate (TPA) were added, and cells were cultured for a further 26 h in the presence of doxycycline and TPA. The cells were then fixed with 4% paraformaldehyde, permeabilized with 0.2% Triton X-100, and labeled with anti-RNA polymerase II mouse monoclonal antibody (Millipore-Sigma, clone CTD4H8, 1:100) and anti-FLAG rabbit polyclonal antibody (Sigma, F7425, 1:100), followed by Alexa 555-anti mouse IgG and Alexa 647-anti rabbit IgG (Thermo Fisher). Nuclei were counterstained with 1μg/ml of Hoechst 33342 (Thermo Fisher). The labeled cells were observed

with a Keyence BZ-X710 fluorescence microscope (Keyence, Osaka, Japan) with standard DAPI, GFP, TRITC and Cy5 filter sets (Chroma Technologies, Bellows Falls, VT, USA).

DNA-FISH combining with LANA immunostaining. BCBL-1 cells were washed with PBS twice and spotted on coverslips and fixed with 4% formaldehyde-PBS for 10 min at RT. Cells were subsequently incubated with PBS with 100 mM glycine to quench residual formaldehyde. Cells were treated with 0.1 % Triton-X and 0.05% SDS with RNaseA for 15 min at 37°C. Coverslips were washed with PBS twice, 70% Ethanol once, 85% Ethanol once, and 100% Ethanol once and dry completely. Cy5-conjugated CENPB-Cy5 PNA probe was obtained from a commercial source (PNA Bio, Thousand Oaks, CA, USA) and coverslips were incubated with hybridization buffer (200 nM PNA probe, 2xSSC, 20% Dextran, 1 mg/mL yeast tRNA, 60% formamide) on glass slide on heat plate at 57°C for 10 min and then 37°C for another 30 min. Utilization of PNA probes allowed us to avoid harsh conditions often used for DNA hybridization, which we found unacceptable noise and it was not readily compatible with subsequent protein staining. Coverslip was washed with warmed 0.1% tween 20 PBS (57°C) in 6 well plates for 3 times and PBS once. Primary antibody was incubated for LANA staining in PBS with 1 mg/mL yeast tRNA. After washing with PBS 3 times, secondly antibody (Alexa 488-anti Rat IgG) was incubated in PBS with yeast tRNA for 1 h at 37°C. After washing three times with PBS, the coverslips were mounted with antifade reagent (Thermo Fisher) and pictures were taken with Keyence microscope as described above.

3D Fluorescence microscopy and quantitative image analysis. Widefield fluorescence imaging was performed on an automated inverted microscope (Keyence BZ-X710) equipped with live-cell chamber. Where applicable: green, red, blue, and far-red fluorescence were

measured using GFP, TRITC, DAPI, and Cy5 filter sets (Chroma Technologies, VT), respectively. Keyence onboard software was used to convert image data to greyscale, TIF format and uniformly process image stacks for background correction, image deconvolution and contrast enhancement.

Image data were subsequently exported to Volocity® Multi-Dimensional Imaging Platform (Quorum Technologies Inc., Ontario, Canada) for 3D visualization and quantitative analysis. For image-based cytometry measurements, individual cells within the field of view were automatically identified and fluorescent signal intensities in the various color channels tabulated. To analyze aggregate formation in the nucleus, image fields were cropped to individual cells – and aggregates were identified based on threshold fluorescence intensities and minimum size characteristics. Statistical information on aggregate size, fluorescence, and colocalization with other measurable aggregates were automatically collected using the Volocity Quantitation toolset. For qualitative assessment of LANA and centromere oligo probes for DNA-FISH, 3D image stacks were similarly compiled and displayed in 3D opacity mode using Volocity Visualization toolset.

Lentivirus production and transduction. For CHD4 knock-down experiments, the following shRNA sequences from Sigma pLKO.1 shRNA libraries were used: shRNA#1: CCTTACTAGAATTGGTGTTAT; shRNA#2: GCTGACACAGTTATTATCTAT. Lentiviruses from CHD4-expressing lentivector, CHD4-targeting shRNAs, and non-targeting scramble shRNA (Addgene, #1864) were produced in 293T cells. The vectors were co-transfected with psPAX2 (Addgene, #12260) and pMD2.G (Addgene, #12259) into 293T cells using polyethylenimine (PEI). Supernatants were collected 48 and 72 h post-transfection. Cells were infected with lentiviruses in the presence of 8 μg/ml polybrene. Cells were incubated for 72 h to allow for

protein expression or knock-down. To generate CHD4 knock-down cell lines, 293T cells were infected with lentivirus for 24 h. Cells were then cultured in selection medium containing 1 μg/ml puromycin for two weeks to obtain stable cell lines.

Preparation of purified KSHV and *de novo* infection. *i*SLK.219 cells latently infected with recombinant KSHV were cultured in eight to ten 150 mm culture dishes until 80% confluent. For reactivation, cultures were re-fed with complete DMEM (without selection drugs) containing 0.3 mM sodium butyrate and 1 μg/ml doxycycline. The cells were further cultured for 4 days in the presence of sodium butyrate and doxycycline. The culture supernatant was centrifuged at 300 x g for 10 min, and then passed through a 0.8-μm filter to remove cellular debris, and then viral particles were concentrated by ultracentrifugation at 25,000 rpm for 2 hrs at 4 °C with Beckman SW28 rotor. Viral precipitate was resuspended in 500 μl of DMEM along with the residual ~500 μl media in the centrifuge tube and stored at –80 °C until use. For *de novo* infection, CHD4-KD 293T cells were seeded at 4 x 10^5 cells/well in 6-well plates. After overnight culture, viral stocks were added to the cultures and the infection was allowed to proceed for 24 h. Cells were monitored for GFP expression and RNA was purified using RNeasy Mini Kit (Qiagen, Venlo, Netherlands).

Reporter assays. Reporter assays were performed as described previously (75). A luciferase reporter system (Promega) was used to quantify ORF16 promoter activity. Briefly, 500 ng of reporter plasmid containing from PAN RNA promoter to in front of translational initiation site of ORF16 was cloned into pGL3 basic vector and used as a reporter. K-Rta responsive elements, insertion of poly(A) sites, and deletion of ENE elements were prepared with mutagenesis. Each reporter construct was transfected along with K-Rta and/or ORF57 expression plasmids with a ratio of 4:1 (reporter 4 and effector 1) in 293 cells. Two days after transfection, the cells were

washed with PBS and incubated with 250 µl passive lysis buffer provided by the manufacturer. At least three independent assays were carried out for each experiment.

Purification of recombinant protein. *Spodoptera frugiperda* Sf9 cells (Millipore) were maintained in Ex-Cell 420 medium (Sigma), and recombinant baculoviruses were generated with BAB to bac system as previously described (76, 77). Transfer plasmid, pFAST-BAC1 vector was modified by inserting Flag tag at N-terminus, and CHD4, ORF57, p65, and Luciferase cDNAs were cloned into CpoI (RsrII) site. Recombinant baculovirus bacmid DNA was transfected into Sf9 cells by using polyethylenimine (Sigma), and recombinant viruses were subsequently amplified twice. Expression of recombinant proteins was confirmed by immunoblotting with anti-Flag monoclonal antibody (Sigma). Large-scale cultures of Sf9 cells (100 ml) were infected with recombinant baculovirus at a multiplicity of infection (MOI) of 0.1 to 1.0, and cells were harvested 48 h after infection. Recombinant proteins were purified as described previously (78). The purity and amount of protein were measured by SDS-PAGE and Coomassie blue staining, using bovine serum albumin (BSA) as a standard.

***In vitro* interaction assays.** Purified full-length Flag-CHD4 and/or Flag-LANA protein was mixed in the binding buffer (20 mM HEPES [pH 7.9], 150 mM NaCl, 1 mM EDTA, 4 mM $MgCl_2$, 1 mM dithiothreitol, 0.02% NP-40, 10% glycerol supplemented with 1 mg/ml BSA, and 1× protease inhibitor cocktail for 30 min at 4°C. After 10-fold dilution with binding buffer, the antibody was incubated at 1:100 dilution and incubated another 1 h at 4°C to form immunocomplex. The immunocomplex was captured with 10 µL of protein A/G magnetic beads mixture. Beads were washed three times with binding buffer and subjected to SDS-PAGE after eluting proteins in sample buffer. The interaction was probed by immunoblotting with either anti-CHD4 (Cell signaling) or anti-LANA (Sigma).

For RNA or DNA pull-down, final 100 nM of each biotinylated RNA fragment (e.g. 1.3 μg for full length PAN RNA transcribed in vitro or 2.6 μg of dsDNA generated by PCR with 5'-biotinylated primer) and purified proteins (100 nM in final) was incubated in RNA-binding buffer (50 mM HEPES [pH 7.9], 150 mM KCl, 5% glycerol, 0.02% NP-40, 100 μg/mL yeast tRNA, 2 mM $MgCl_2$) supplemented with RNAse inhibitor in 40 μL reaction for 1 h at 4°C. Biotinylated RNA or dsDNA was captured with streptavidin conjugated magnetic beads and proteins were eluted in sample buffer. Precipitated proteins were visualized by immunoblotting with either anti-Flag antibody (Sigma) or specific antibody as indicated in figure legends (LANA or CHD4).

Quantification of viral replication. siRNA targeting selected cellular genes were transfected in *i*SLK.219 cells for 48h followed by KSHV reactivation by doxycycline (1 μg/ml). After 24h, the RFP fluorescence intensity was quantified using ImageJ software. The RFP signal intensity was normalized relative to that in cells transfected with non-targeting siRNA (NTC).

Construction of mini-Turbo KSHV BAC16. Recombinant KSHV was prepared by following a protocol for En passant mutagenesis with a two-step markerless red recombination technique (79). Briefly, the codon optimized the mini-TurboID coding sequence (Table 1), which also encodes the 3x Flag tag was first cloned into a pBS SK vector (Thermo Fisher, Waltham, MA USA). The pEPkan-S plasmid was used as a source of the kanamycin cassette, which includes the I-SecI restriction enzyme site at the 5'-end of the kanamycin coding region (79). Kanamycin cassette was amplified with primer pairs listed in Table 1, and cloned into the mini-TurboID coding region at a unique restriction enzyme site. The resulting plasmid was used as a template for another round of PCR to prepare a transfer DNA fragment for markerless recombination with BAC16 (80). Recombinant BAC clones with insertion and also deletion of the kanamycin cassette in the BAC16 genome were confirmed by colony PCR with appropriate primer pairs.

Recombination junctions and adjacent genomic regions were amplified by PCR and the resulting PCR products were directly sequenced with the same primers to confirm in-frame insertion of the mini-TurboID cassette into the BAC DNA. The resulting recombinant BAC was confirmed by restriction enzyme digestions (*Hind*III and *Bgl*II), if there were any large DNA deletions. Two independent BAC clones were generated for each mini-TurboID tagged recombinant virus as biological replicates, and used one of the clone for protein ID. Entire BAC DNAs were subsequently sequenced. Insertion of PAN RNA MRE mutations was performed by using primer pairs that encode the intended mutations. Primer pairs were used to amplify the kanamycin cassette and recombination, deletion of kanamycin cassette, confirmation of mutations was performed as described above. Similarly, three stop codons were introduced into the first ORF57 exon with BAC recombination along with HA epitopes were inserted as an N-terminal fusion.

Western blotting. Cells were lysed in IP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% NP-40, 1 mM EDTA, 5% glycerol) containing protease inhibitors (Roche, Basel, Switzerland). Total cell lysates (25 μg) were boiled in SDS-PAGE loading buffer and subjected to SDS-PAGE and subsequently transferred to a polyvinylidene fluoride membrane (Millipore-Sigma, St. Louis, MO, USA) using a semidry transfer apparatus (Bio-Rad, Hercules, CA, USA). The streptavidin-HRP conjugate was used at 1:3000 dilution. Final dilution of the primary antibody was 1:5,000 for anti-K-Rta rabbit serum, 1 μg/ml of anti-K8α (Santa Cruz, Santa Cruz, CA, USA), 1 μg/mL of anti-ORF57 mouse monoclonal antibody (Santa Cruz, Santa Cruz, CA, USA), 1 μg/ml of anti-K8.1 mouse monoclonal (Santa Cruz, Santa Cruz, CA, USA), and 1:5,000 for anti-β-actin mouse monoclonal (Millipore-Sigma, St. Louis, MO, USA). Washing membranes and secondary antibody incubations were performed as described previously (77).

Quantification of viral copy number. Two hundred microliter of cell culture supernatant was treated with 12 μg/ml of DNase I for 15 min at room temperature to degrade uncapsidated DNA. This reaction was stopped by the addition of EDTA to 5 mM followed by heating at 70°C for 15 min. Viral genomic DNA was purified using QIAamp DNA Mini Kit according to the manufacturer's protocol, and eluted in 100 μl of buffer AE. Four microliters of the eluate were used for real-time qPCR to determine viral copy number, as described previously (77).

Real-time RT-PCR. Total RNA was isolated using the Quick-RNA miniprep kit (Zymo Research, Irvine, CA, USA). First-strand cDNA was synthesized using High Capacity cDNA Reverse Transcription Kit (Thermo Fisher, Waltham, MA USA). Gene expression was analyzed by realtime qPCR using specific primers for KSHV ORFs designed by Fakhari and Dittmer (81). We used 18S ribosomal RNA or GAPDH as an internal standard to normalize viral gene expression.

Affinity purification of biotinylated proteins. The affinity purification was done with streptavidin-coated magnetic beads (Thermo-Fisher). Briefly, 150 μl magnetic beads/sample were pre-washed with RIPA lysis buffer (150 mM NaCl, 5 mM EDTA (pH 8), 50 mM Tris (pH 8), 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) 3 times. Total 3 mg of whole cell lysate was incubated with pre-washed streptavidin beads at room temperature for 1h with rotation. The beads were collected using a magnetic stand and washed three times with wash buffer according to the manufacturer's protocol. Finally, beads were resuspended in 200 μl of wash buffer and sent to UC Davis Proteomics core for on bead digestion and LC-MS/MS analysis.

MS sample preparation. Protein samples on magnetic beads were washed four times with 200 μl of 50mM ammonium bicarbonate (AMBIC) with a twenty-minute shake time at 4°C in between

each wash. Roughly 2.5 μg of trypsin was added to the bead and AMBIC and the samples were digested overnight at 800 rpm shake speed. After overnight digestion, the supernatant was removed, and the beads were washed once with enough 50 mM ammonium bicarbonate to cover. After 20 minutes at a gentle shake the wash is removed and combined with the initial supernatant. The peptide extracts are reduced in volume by vacuum centrifugation and a small portion of the extract is used for fluorometric peptide quantification (Thermo Scientific Pierce). One microgram of sample based on the fluorometric peptide assay was loaded for each LC-MS analysis.

Digested peptides were analyzed by LC-MS/MS on a Thermo Scientific Q Exactive Orbitrap Mass spectrometer in conjunction with Proxeon Easy-nLC II HPLC (Thermo Scientific) and Proxeon nanospray source. The digested peptides were loaded a 100 micron x 25 mm Magic C18 100Å 5U reverse phase trap where they were desalted online before being separated using a 75-micron x 150 mm Magic C18 200Å 3U reverse-phase column. Peptides were eluted using a 60-minute gradient with a flow rate of 300 μl/min. An MS survey scan was obtained for the m/z range 300-1600, MS/MS spectra were acquired using a top 15 method, where the top 15 ions in the MS spectra were subjected to HCD (High Energy Collisional Dissociation). An isolation mass window of 2.0 m/z was for the precursor ion selection, and normalized collision energy of 27% was used for fragmentation. A fifteen second duration was used for the dynamic exclusion.

MS/MS analysis. Tandem mass spectra were extracted and charge state deconvoluted by Proteome Discoverer (Thermo Scientific). All MS/MS samples were analyzed using X! Tandem (The GPM, thegpm.org; version X! Tandem Alanine (2017.2.1.4)). X! Tandem was set up to search the Human and Kaposi Sarcoma Herpes virus database (149182 entries) assuming the

digestion enzyme trypsin. X! Tandem was searched with a fragment ion mass tolerance of 20 PPM and a parent ion tolerance of 20 PPM. Carbamidomethyl of cysteine and selenocysteine was specified in X! Tandem as a fixed modification. Glu->pyro-Glu of the N-terminus, ammonia-loss of the N-terminus, gln->pyro-Glu of the N-terminus, deamidated of asparagine and glutamine, oxidation of methionine and tryptophan and dioxidation of methionine and tryptophan were specified in X! Tandem as variable modifications.

Scaffold (version Scaffold_4.8.4, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 98.0% probability by the Scaffold Local FDR algorithm. Peptide identifications were also required to exceed specific database search engine thresholds. Protein identifications were accepted if they could be established at greater than 5.0% probability to achieve an FDR less than 5.0% and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (82). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

Pathway analysis. The proteins identified to be interacting with ORF57 were used for Gene ontology analysis. The top gene ontology processes were enriched by the Metascape web-based platform, and the Metascape software was used for gene ontology analysis (83).

Statistical analysis. Results are shown as mean ± SD from at least three independent experiments. Data were analyzed using unpaired Student's t test, or ANOVA followed by Tukey’s HSD test. A value of $p<0.05$ was considered statistically significant.

Code availability. All custom codes used to analyse data and generate figures are available upon reasonable request.

Data availability. Data can be accessed via the Source Data files and Supplementary tables. The data discussed in this publication have been deposited in NCBI's GEO Database under unified Super Series GSE163695 with subseries GSE163098 for CUT&RUN and GSE163694 for CHi-C.

Acknowledgement. This research was supported by public health grants from the National Cancer Institute (CA225266, CA232845), the National Institute of Dental and Craniofacial (DE025985), and the National Institute of Allergy and Infectious Disease (AI147207) to Y.I. The Genomics Shared Resource is supported by the UC Davis Comprehensive Cancer Center Support Grant (CCSG) awarded by the National Cancer Institute (NCI P30CA093373).

Author Contributions.

A.K., Y.L., M.C., and Y.I. designed and performed experiments, analysed data and prepared figures. Y.Y., C.C., R.R.D., K.N., and A.M. run genomic pipelines, performed statistical analyses, and visualized genomic data sets. M.S. prepared enzyme digested protein samples and performed proteomics data analyses. R.R.D. and C.G.T. prepared sequencing libraries and performed initial bioinformatics analyses. K.I.N and R.R.D. prepared single cell transcriptomics libraries, and A.M. wrote custom codes and analyzed single-cell sequence data sets. A.K., Y.L., K.W., C.I., and Y.I. prepared BAC stable cell lines, and performed molecular and biochemical studies. YI generated recombinant BACs. F.C. and K.I.N. prepared slides and performed imaging analyses. V.M., M.S., and M.Z. contributed key reagents. A.K, Y.L., Y.Y., C.C., M.S., F.C., R.R.D., C.G.T., K.N., K.I.N.,

and Y.I. analyzed corresponding data sets. Y.I. wrote a manuscript and all authors edited the manuscript.

Competing financial interests.

The authors have declared that no conflicts of interests exist.

Materials & Correspondence.

Correspondence and request for materials should be addressed to M.C. (e-mail: mcampbell@ucdavis.edu) or Y.I. (email: yizumiya@ucdavis.edu).

Figure legends

Fig. 1 KSHV RNA-binding protein functions in genomic looping. (a) Immunoblotting. A recombinant KSHV harboring stop codons in ORF57 open reading frame was generated (ORF57-Stop KSHV) with BAC16. Knock-out of ORF57 protein tagged with a HA epitope was confirmed and the expression of other viral proteins was also confirmed with the indicated antibodies. **(b) Frequencies of genomic looping.** Frequencies of genomic interactions within KSHV genomes in ORF57-Wt (left) or ORF57-Stop (right) KSHV infected cells were plotted by TADbit as heatmap (Bin: 0.5 kb). Insulation score and border strength of TADs in KSHV genomes were also analyzed and plotted in the bottom of panels. **(c) Gene-density plots and relative changes in genomic looping by ORF57 knock-out.** Gene densities, which indicates frequencies of neighboring to other genes, are shown (Bin: 2 kb). Differences in the genomic loops between ORF57-Wt and ORF57-Stop in reactivated cells were calculated by subtracting from relative abundances of genomic loop in ORF57-Stop and shown as heatmap (bottom panel). The heatmap was plotted with Juicebox. **(d) 3C analyses.** The PAN RNA genomic region was used as anchor sites to perform 3C with qPCR. KSHV BAC DNA was used as control. **(e) RT-qPCR.** Expression of selected KSHV genes from different kinetics classes was measured by RT-qPCR and relative abundance was visualized as a heatmap. The amount of transcripts at 0 hr time point was set as the baseline (i.e. 1). **(f) Viral replication.** KSHV virion production was measured by encapsidated KSHV DNAs in culture supernatant after 96 hr reactivation with K-Rta induction with TPA. **** $P<0.0001$.

Fig. 2 Induction of RNA polymerase II aggregation on KSHV genomes. (a) RNA-FISH and IFA. Recombinant KSHV infected *i*SLK cells were induced for reactivation by induction of K-Rta

expression and TPA treatment for 28 hours. PAN RNA, RNA pol II, and ORF57 protein were visualized by either specific oligo-probes or antibodies (anti-HA tag for ORF57). Volumes of aggregates and intensity of signals were measured with Velocity software. **(b) RNA pol II recruitment.** CUT&RUN was performed with an anti-RNA pol II antibody. Occupancy sites were visualized with the Integrative Genomics Viewer (IGV) and screenshots are presented that depict RNA pol II recruitment before (latency; top panel) and 28 hours after reactivation (bottom panel). Blue lines; ORF57-wt cells, Pink line; ORF57-Stop cells. **(c) Distal promoter activation by PAN RNA expression**. A schematic diagram of the reporter constructs is shown. Fold activation over vector control transfected is provided in bar charts. Total amount of plasmids transfected was normalized by supplementing empty vector and assays were performed with HEK293 cells in triplicated wells. Data are means ± standard deviation of triplicate samples from a single experiment and are representative of three independent experiments (c).

Fig. 3 Proximity biotin labeling with mini-TurboID-ORF57 to identify PAN RNA binding proteins. (a) A schematic diagram of recombinant KSHVs. Mini-TurboID is fused to the N-terminus of ORF57 open reading frame and the ORF57 binding site (MRE) was mutated as shown. **(b) KSHV RT-qPCR arrays.** The genome-wide expression of KSHV genes and lncRNAs was examined by PCR arrays and fold activation over 0 hr time point was plotted as heat maps. **(c) Immunoblotting.** Recombinant KSHV infected *i*SLK cells were either induced reactivation with doxycycline + TPA for 24 hours, protein was probed with either anti-Flag antibody (for mini-TurboID-ORF57) or Streptavidin-HRP (for biotinylated proteins). **(d) Venn diagram.** Venn diagram indicates the number of identified proteins with a P value less than 0.05. Each protein ID was performed in three biological replicates. **(e) PAN RNA-mediated interacting proteins.**

Selected protein names were listed. The entire list of proteins and peptide counts is also presented in Extended Data Table 1. **(f) Gene Ontology analyses of PAN RNA-interacting proteins.** Non-redundant top 5 biological pathways are presented.

Fig. 4 CHD4 is an RNA and DNA binding protein and a strong silencer of KSHV gene expression. (a) siRNA screening with KSHV reactivation as readout. Recombinant KSHV reactivation, which encodes RFP under control of the PAN RNA promoter, was used to screen the effects of KSHV reactivation. RFP signal intensity was measured with Image J, and the GFP signal was used as internal controls. Three randomly selected fields in the middle of each wells were quantified and the average intensity was plotted. Representative pictures were also shown in right panels. siPOLR2A was used as a positive experimental control. **(b) CHD4 overexpression and complementation.** Encapsidated KSHV genomes in the culture supernatant were measured by qPCR to assess effects of CHD4 expression on KSHV replication. CHD4 was over expressed, knocked-down, or knocked-down then compensated by transfection of mouse CHD4 cDNA or transduction of shRNA. **(c) CHD4 ATPase activity is important for prevention of nuclear aggregate formation.** Flag tagged mouse CHD4 or mutant was transfected and stained with anti-Flag. Transcriptional factory formation was assessed by staining RNA polymerase II. The percentage of aggregate formation in Flag positive cells was measured and plotted. Three biological replicates and counted 50 cells were counted in each of three biological replicates. **(d) Recombinant proteins and synthesized PAN RNAs.** Recombinant proteins were expressed with baculoviruses and purified with Flag-agarose beads. Coomassie staining of Flag-ORF57 and Flag-hCHD4 are shown. **(e) CHD4 binds RNAs.** RNA pull down was performed with indicated biotinylated RNAs and visualized by immunoblotting

with anti-Flag antibody. Beads alone was used for background control. ORF57 protein did not compete or enhance CHD4 RNA binding (right panel). **(f) CHD4 binding sites on the KSHV genome**. CUT&RUN was performed with BC-1 by using indicated antibodies and sequence reads were mapped to KSHV genome. Snapshot of IGV viewer and KSHV genome map were shown. Results of similar experiments with BCBL-1 and BC3 were presented in Expanded Data Figure 4. **(g) PAN RNA competes with CHD4 binding on PAN RNA genomic region.** Pull-down analyses with biotinylated ssRNA or dsDNA was performed. CHD4 (100 nM) was incubated with 100 nM of biotinylated RNA or 100 nM of biotinylated dsDNA in 40 uL binding buffer. Increased amount of non-biotinylated PAN RNA at 1:1, 1:10, and 1:20 (dsDNA vs. ssRNA) was also incubated and precipitated CHD4 protein in the pulldown was probed with anti-Flag antibody. Flow-through and 10% of reaction before pull-down were used as control. An equal amount of dsDNA on beads was also confirmed by qPCR. Data are means ±SEM of triplicate samples from a single experiment and are representative of three independent experiments (**a, b)**.

Fig. 5 KSHV episomes colocalize with ChAHP complex on host chromatins. (a) KSHV episome docking sites on host cell chromosomes. CHi-C chimeric DNA ligation products composed of sequences derived from the KSHV and human genomes were mapped in three PEL cells. For BCBL-1 cells, three biological replicates were performed with nearly identical results (similarity 0.95). Reads counts of sum of 10 kb Bin were plotted. **(b) Zoomed selected chromosome view.** Selected chromosome views were also depicted, which shows enrichment of sequence reads near centromeres. Extended panels were also presented in Extended Data Figure 4. **(c) Association of KSHV episomes docking sites with LANA and CHD4**

recruitments. CUT&RUN sequence reads with anti-LANA or anti-CHD4 were mapped and visualized with IGV. A snapshot of the IGV view is presented. **(d) Jaccard statistics and Venn diagram.** CUT&RUN peaks spanning more than 1 kb were isolated and analyzed co-localization on host chromosome with Jaccard statistics. Venn diagram showed the intersection of genomic regions with the length more than 1,000 bp of 4 different proteins and the Jaccard statistic was used to compare similarity of positions of those genes. **(e) A detachment of KSHV episomes by reactivation.** Relative C-HiC reads were visualized by Juicebox as similar to Fig. 1c. **(f) KSHV episomes locates near the centromere.** DNA-FISH was used to visualize location of centromere and KSHV episomes were located with anti-LANA antibody with IFA. Two signals were frequently localized next each other. The 3D view was presented in Extended Data Figure 8. **(g) mini-TurboID LANA.** Recombinant KSHV, which harbors mini-TurboID fused LANA was generated and infected in iSLK cells. Protein localizing in the proximity of LANA were identified with a mass spectrometry. Selected proteins with p value less than 0.05 were listed. Each protein ID was performed in three biological replicates. T-test was used for calculating the p value. Bold case; ChAHP components. Underlined; previously identified proteins as LANA interacting proteins. **(h) LANA interacts with CHD4 *in vitro*.** Purified LANA and CHD4 were prepared from recombinant baculovirus infected Sf9 cells. Coomassie staining is shown. The equal molar amount (100 nM) of two proteins were incubated in binding buffer and immunoprecipitated with the specific antibodies of one another. The 10% of reaction before immunoprecipitation was shown as controls.

Fig. 6 CHD4 is important for establishing latency. (a) CHD4 knockdown and KSHV lytic gene expression during *de novo* infection. Immunoblot analysis of CHD4 protein and RT-

qPCR analysis of CHD4 mRNA expression in scramble shRNA- (shScrm) or shCHD4-transduced stable 293T cell lines are shown. 293T cells stably transduced with shScrm or shCHD4 were infected with rKSHV.219 for 24 h. Total RNA was harvested at the indicated time points. K-Rta and ORF25 expression was measured by RT-qPCR and normalized to GAPDH. * P<0.05, ** P<0.01 (comparing with shScramble control). **(b) KSHV latency model.** KSHV episomes tether at borders of host TADs with ChAHP complex. Stimulation of K-Rta, in part by decreased CHD4 expression, triggered robust PAN RNA expression, which compete CHD4 DNA binding to facilitate KSHV reactivation. Leaving CHD4 enriched nuclear environment by detaching from host chromatin further enhance KSHV reactivation.

Reference

1. **Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, Knowles DM, Moore PS.** 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science **266:**1865-1869.
2. **Ganem D.** 2010. KSHV and the pathogenesis of Kaposi sarcoma: listening to human biology and medicine. J Clin Invest **120:**939-949.
3. **Mesri EA, Cesarman E, Boshoff C.** 2010. Kaposi's sarcoma and its associated herpesvirus. Nat Rev Cancer **10:**707-719.
4. **Schulz TF.** 2006. The pleiotropic effects of Kaposi's sarcoma herpesvirus. J Pathol **208:**187-198.
5. **Wong EL, Damania B.** 2005. Linking KSHV to human cancer. Curr Oncol Rep **7:**349-356.
6. **Cesarman E, Moore PS, Rao PH, Inghirami G, Knowles DM, Chang Y.** 1995. In vitro establishment and characterization of two acquired immunodeficiency syndrome-related lymphoma cell lines (BC-1 and BC-2) containing Kaposi's sarcoma-associated herpesvirus-like (KSHV) DNA sequences. Blood **86:**2708-2714.
7. **Cesarman E, Knowles DM.** 1999. The role of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) in lymphoproliferative diseases. Semin Cancer Biol **9:**165-174.
8. **Nakamura H, Lu M, Gwack Y, Souvlis J, Zeichner SL, Jung JU.** 2003. Global changes in Kaposi's sarcoma-associated virus gene expression patterns following expression of a tetracycline-inducible Rta transactivator. J Virol **77:**4205-4220.
9. **Lu M, Suen J, Frias C, Pfeiffer R, Tsai MH, Chuang E, Zeichner SL.** 2004. Dissection of the Kaposi's sarcoma-associated herpesvirus gene expression program by using the viral DNA replication inhibitor cidofovir. J Virol **78:**13637-13652.
10. **Schoenfelder S, Fraser P.** 2019. Long-range enhancer-promoter contacts in gene expression control. Nat Rev Genet **20:**437-455.
11. **Campbell M, Watanabe T, Nakano K, Davis RR, Lyu Y, Tepper CG, Durbin-Johnson B, Fujimuro M, Izumiya Y.** 2018. KSHV episomes reveal dynamic chromatin loop formation with domain-specific gene regulation. Nat Commun **9:**49.
12. **Chen CP, Lyu Y, Chuang F, Nakano K, Izumiya C, Jin D, Campbell M, Izumiya Y.** 2017. Kaposi's Sarcoma-Associated Herpesvirus Hijacks RNA Polymerase II To Create a Viral Transcriptional Factory. J Virol **91**.
13. **Sun R, Lin SF, Gradoville L, Miller G.** 1996. Polyadenylylated nuclear RNA encoded by Kaposi sarcoma-associated herpesvirus. Proc Natl Acad Sci U S A **93:**11883-11888.
14. **Conrad NK, Steitz JA.** 2005. A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. EMBO J **24:**1831-1841.
15. **Massimelli MJ, Kang JG, Majerciak V, Le SY, Liewehr DJ, Steinberg SM, Zheng ZM.** 2011. Stability of a long noncoding viral RNA depends on a 9-nt core element at the RNA 5' end to interact with viral ORF57 and cellular PABPC1. Int J Biol Sci **7:**1145-1160.
16. **Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA.** 2010. Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. Science **330:**1244-1247.
17. **Massimelli MJ, Majerciak V, Kruhlak M, Zheng ZM.** 2013. Interplay between polyadenylate-binding protein 1 and Kaposi's sarcoma-associated herpesvirus ORF57 in accumulation of polyadenylated nuclear RNA, a viral long noncoding RNA. J Virol **87:**243-256.
18. **Rossetto CC, Tarrant-Elorza M, Verma S, Purushothaman P, Pari GS.** 2013. Regulation of viral and cellular gene expression by Kaposi's sarcoma-associated herpesvirus polyadenylated nuclear RNA. J Virol **87:**5540-5553.
19. **Campbell M, Kim KY, Chang PC, Huerta S, Shevchenko B, Wang DH, Izumiya C, Kung HJ, Izumiya Y.** 2014. A lytic viral long noncoding RNA modulates the function of a latent protein. J Virol **88:**1843-1848.
20. **Rossetto CC, Pari G.** 2012. KSHV PAN RNA associates with demethylases UTX and JMJD3 to activate lytic replication through a physical interaction with the virus genome. PLoS Pathog **8:**e1002680.

21. **Borah S, Darricarrere N, Darnell A, Myoung J, Steitz JA.** 2011. A viral nuclear noncoding RNA binds re-localized poly(A) binding protein and is required for late KSHV gene expression. PLoS Pathog **7:**e1002300.
22. **Majerciak V, Zheng ZM.** 2015. KSHV ORF57, a protein of many faces. Viruses **7:**604-633.
23. **Withers JB, Li ES, Vallery TK, Yario TA, Steitz JA.** 2018. Two herpesviral noncoding PAN RNAs are functionally homologous but do not associate with common chromatin loci. PLoS Pathog **14:**e1007389.
24. **Sharma NR, Majerciak V, Kruhlak MJ, Yu L, Kang JG, Yang A, Gu S, Fritzler MJ, Zheng ZM.** 2019. KSHV RNA-binding protein ORF57 inhibits P-body formation to promote viral multiplication by interaction with Ago2 and GW182. Nucleic Acids Res **47:**9368-9385.
25. **Chujo T, Yamazaki T, Hirose T.** 2016. Architectural RNAs (arcRNAs): A class of long noncoding RNAs that function as the scaffold of nuclear bodies. Biochim Biophys Acta **1859:**139-146.
26. **Quinn JJ, Chang HY.** 2016. Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet **17:**47-62.
27. **Nakagawa S, Yamazaki T, Hirose T.** 2018. Molecular dissection of nuclear paraspeckles: towards understanding the emerging world of the RNP milieu. Open Biol **8**.
28. **Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, Bickle M, Rizk S, Guillen-Boixet J, Franzmann TM, Jahnel M, Marrone L, Chang YT, Sterneckert J, Tomancak P, Hyman AA, Alberti S.** 2018. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science **360:**918-921.
29. **Yamazaki T, Souquere S, Chujo T, Kobelke S, Chong YS, Fox AH, Bond CS, Nakagawa S, Pierron G, Hirose T.** 2018. Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. Mol Cell **70:**1038-1053 e1037.
30. **Hansen AS, Hsieh TS, Cattoglio C, Pustova I, Saldana-Meyer R, Reinberg D, Darzacq X, Tjian R.** 2019. Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. Mol Cell **76:**395-411 e313.
31. **Saldana-Meyer R, Rodriguez-Hernaez J, Escobar T, Nishana M, Jacome-Lopez K, Nora EP, Bruneau BG, Tsirigos A, Furlan-Magaril M, Skok J, Reinberg D.** 2019. RNA Interactions Are Essential for CTCF-Mediated Genome Organization. Mol Cell **76:**412-422 e415.
32. **Majerciak V, Pripuzova N, McCoy JP, Gao SJ, Zheng ZM.** 2007. Targeted disruption of Kaposi's sarcoma-associated herpesvirus ORF57 in the viral genome is detrimental for the expression of ORF59, K8alpha, and K8.1 and the production of infectious virus. J Virol **81:**1062-1071.
33. **Xiao R, Chen JY, Liang Z, Luo D, Chen G, Lu ZJ, Chen Y, Zhou B, Li H, Du X, Yang Y, San M, Wei X, Liu W, Lecuyer E, Graveley BR, Yeo GW, Burge CB, Zhang MQ, Zhou Y, Fu XD.** 2019. Pervasive Chromatin-RNA Binding Protein Interactions Enable RNA-Based Regulation of Transcription. Cell **178:**107-121 e118.
34. **Rieder D, Trajanoski Z, McNally JG.** 2012. Transcription factories. Front Genet **3:**221.
35. **Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY.** 2018. Efficient proximity labeling in living cells and organisms with TurboID. Nat Biotechnol **36:**880-887.
36. **Weiss K, Lazar HP, Kurolap A, Martinez AF, Paperna T, Cohen L, Smeland MF, Whalen S, Heide S, Keren B, Terhal P, Irving M, Takaku M, Roberts JD, Petrovich RM, Schrier Vergano SA, Kenney A, Hove H, DeChene E, Quinonez SC, Colin E, Ziegler A, Rumple M, Jain M, Monteil D, Roeder ER, Nugent K, van Haeringen A, Gambello M, Santani A, Medne L, Krock B, Skraban CM, Zackai EH, Dubbs HA, Smol T, Ghoumid J, Parker MJ, Wright M, Turnpenny P, Clayton-Smith J, Metcalfe K, Kurumizaka H, Gelb BD, Baris Feldman H, Campeau PM, Muenke M, Wade PA, Lachlan K.** 2020. The CHD4-related syndrome: a comprehensive investigation of the clinical spectrum, genotype-phenotype correlations, and molecular basis. Genet Med **22:**389-397.
37. **Farnung L, Ochmann M, Cramer P.** 2020. Nucleosome-CHD4 chromatin remodeler structure maps human disease mutations. Elife **9**.
38. **Morra R, Lee BM, Shaw H, Tuma R, Mancini EJ.** 2012. Concerted action of the PHD, chromo and motor domains regulates the human chromatin remodelling ATPase CHD4. FEBS Lett **586:**2513-2521.

39. **Viejo-Borbolla A, Ottinger M, Bruning E, Burger A, Konig R, Kati E, Sheldon JA, Schulz TF.** 2005. Brd2/RING3 interacts with a chromatin-binding domain in the Kaposi's Sarcoma-associated herpesvirus latency-associated nuclear antigen 1 (LANA-1) that is required for multiple functions of LANA-1. J Virol **79:**13618-13629.
40. **De Leon Vazquez E, Juillard F, Rosner B, Kaye KM.** 2014. A short sequence immediately upstream of the internal repeat elements is critical for KSHV LANA mediated DNA replication and impacts episome persistence. Virology **448:**344-355.
41. **Barbera AJ, Chodaparambil JV, Kelley-Clarke B, Joukov V, Walter JC, Luger K, Kaye KM.** 2006. The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. Science **311:**856-861.
42. **Arnold PR, Wells AD, Li XC.** 2019. Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. Front Cell Dev Biol **7:**377.
43. **Nair SJ, Yang L, Meluzzi D, Oh S, Yang F, Friedman MJ, Wang S, Suter T, Alshareedah I, Gamliel A, Ma Q, Zhang J, Hu Y, Tan Y, Ohgi KA, Jayani RS, Banerjee PR, Aggarwal AK, Rosenfeld MG.** 2019. Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. Nat Struct Mol Biol **26:**193-203.
44. **Isoda T, Moore AJ, He Z, Chandra V, Aida M, Denholtz M, Piet van Hamburg J, Fisch KM, Chang AN, Fahl SP, Wiest DL, Murre C.** 2017. Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. Cell **171:**103-119 e118.
45. **Coe BP, Stessman HAF, Sulovari A, Geisheker MR, Bakken TE, Lake AM, Dougherty JD, Lein ES, Hormozdiari F, Bernier RA, Eichler EE.** 2019. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat Genet **51:**106-116.
46. **Kovac K, Sauer A, Macinkovic I, Awe S, Finkernagel F, Hoffmeister H, Fuchs A, Muller R, Rathke C, Langst G, Brehm A.** 2018. Tumour-associated missense mutations in the dMi-2 ATPase alters nucleosome remodelling properties in a mutation-specific manner. Nat Commun **9:**2112.
47. **Zhao H, Han Z, Liu X, Gu J, Tang F, Wei G, Jin Y.** 2017. The chromatin remodeler Chd4 maintains embryonic stem cell identity by controlling pluripotency- and differentiation-associated genes. J Biol Chem **292:**8507-8519.
48. **Sifrim A, Hitz MP, Wilsdon A, Breckpot J, Turki SH, Thienpont B, McRae J, Fitzgerald TW, Singh T, Swaminathan GJ, Prigmore E, Rajan D, Abdul-Khaliq H, Banka S, Bauer UM, Bentham J, Berger F, Bhattacharya S, Bu'Lock F, Canham N, Colgiu IG, Cosgrove C, Cox H, Daehnert I, Daly A, Danesh J, Fryer A, Gewillig M, Hobson E, Hoff K, Homfray T, Study I, Kahlert AK, Ketley A, Kramer HH, Lachlan K, Lampe AK, Louw JJ, Manickara AK, Manase D, McCarthy KP, Metcalfe K, Moore C, Newbury-Ecob R, Omer SO, Ouwehand WH, Park SM, Parker MJ, Pickardt T, Pollard MO, et al.** 2016. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. Nat Genet **48:**1060-1065.
49. **Zhao S, Choi M, Overton JD, Bellone S, Roque DM, Cocco E, Guzzo F, English DP, Varughese J, Gasparrini S, Bortolomai I, Buza N, Hui P, Abu-Khalaf M, Ravaggi A, Bignotti E, Bandiera E, Romani C, Todeschini P, Tassi R, Zanotti L, Carrara L, Pecorelli S, Silasi DA, Ratner E, Azodi M, Schwartz PE, Rutherford TJ, Stiegler AL, Mane S, Boggon TJ, Schlessinger J, Lifton RP, Santin AD.** 2013. Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma. Proc Natl Acad Sci U S A **110:**2916-2921.
50. **Le Gallo M, O'Hara AJ, Rudd ML, Urick ME, Hansen NF, O'Neil NJ, Price JC, Zhang S, England BM, Godwin AK, Sgroi DC, Program NIHISCCS, Hieter P, Mullikin JC, Merino MJ, Bell DW.** 2012. Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. Nat Genet **44:**1310-1315.
51. **Goodman JV, Yamada T, Yang Y, Kong L, Wu DY, Zhao G, Gabel HW, Bonni A.** 2020. The chromatin remodeling enzyme Chd4 regulates genome architecture in the mouse brain. Nat Commun **11:**3419.

52. **Marques JG, Gryder BE, Pavlovic B, Chung Y, Ngo QA, Frommelt F, Gstaiger M, Song Y, Benischke K, Laubscher D, Wachtel M, Khan J, Schafer BW.** 2020. NuRD subunit CHD4 regulates super-enhancer accessibility in rhabdomyosarcoma and represents a general tumor dependency. Elife **9**.
53. **Naik NG, Nguyen TH, Roberts L, Fischer LT, Glickman K, Golas G, Papp B, Toth Z.** 2020. Epigenetic factor siRNA screen during primary KSHV infection identifies novel host restriction factors for the lytic cycle of KSHV. PLoS Pathog **16:**e1008268.
54. **Ostapcuk V, Mohn F, Carl SH, Basters A, Hess D, Iesmantavicius V, Lampersberger L, Flemr M, Pandey A, Thoma NH, Betschinger J, Buhler M.** 2018. Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature **557:**739-743.
55. **Kaaij LJT, Mohn F, van der Weide RH, de Wit E, Buhler M.** 2019. The ChAHP Complex Counteracts Chromatin Looping at CTCF Sites that Emerged from SINE Expansions in Mouse. Cell **178:**1437-1451 e1414.
56. **Kim KD, Tanizawa H, De Leo A, Vladimirova O, Kossenkov A, Lu F, Showe LC, Noma KI, Lieberman PM.** 2020. Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus. Nat Commun **11:**877.
57. **Izumiya Y, Izumiya C, Van Geelen A, Wang DH, Lam KS, Luciw PA, Kung HJ.** 2007. Kaposi's sarcoma-associated herpesvirus-encoded protein kinase and its interaction with K-bZIP. J Virol **81:**1072-1082.
58. **Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL.** 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell **159:**1665-1680.
59. **Belaghzal H, Dekker J, Gibcus JH.** 2017. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. Methods **123:**56-65.
60. **Schmitt AD, Hu M, Ren B.** 2016. Genome-wide mapping and analysis of chromosome architecture. Nature Reviews Molecular Cell Biology **17:**743-755.
61. **Schoenfelder S, Javierre B-M, Furlan-Magaril M, Wingett SW, Fraser P.** 2018. Promoter Capture Hi-C: High-resolution, Genome-wide Profiling of Promoter Interactions. Journal of Visualized Experiments : JoVE doi:10.3791/57320.
62. **Jung I, Schmitt A, Diao Y, Lee AJ, Liu T, Yang D, Tan C, Eom J, Chan M, Chee S, Chiang Z, Kim C, Masliah E, Barr CL, Li B, Kuan S, Kim D, Ren B.** 2019. A compendium of promoter-centered long-range chromatin interactions in the human genome. Nat Genet **51:**1442-1449.
63. **Skene PJ, Henikoff S.** 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife **6**.
64. **Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E.** 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol **16:**259.
65. **Serra F, Bau D, Goodstadt M, Castillo D, Filion GJ, Marti-Renom MA.** 2017. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. PLoS Comput Biol **13:**e1005665.
66. **Trussart M, Serra F, Bau D, Junier I, Serrano L, Marti-Renom MA.** 2015. Assessing the limits of restraint-based 3D modeling of genomes and genomic domains. Nucleic Acids Res **43:**3465-3477.
67. **Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, Andrews S.** 2015. HiCUP: pipeline for mapping and processing Hi-C data. F1000Res **4:**1310.
68. **Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL.** 2016. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst **3:**99-101.
69. **Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK.** 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell **38:**576-589.
70. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods **9:**357-359.
71. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS.** 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol **9:**R137.

72. **Freese NH, Norris DC, Loraine AE.** 2016. Integrated genome browser: visual analytics platform for genomics. Bioinformatics **32:**2089-2095.
73. **Khan A, Mathelier A.** 2017. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. BMC Bioinformatics **18:**287.
74. **Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P, Satija R.** 2019. Comprehensive Integration of Single-Cell Data. Cell **177:**1888-1902 e1821.
75. **Ellison TJ, Izumiya Y, Izumiya C, Luciw PA, Kung HJ.** 2009. A comprehensive analysis of recruitment and transactivation potential of K-Rta and K-bZIP during reactivation of Kaposi's sarcoma-associated herpesvirus. Virology **387:**76-88.
76. **Kim KY, Huerta SB, Izumiya C, Wang DH, Martinez A, Shevchenko B, Kung HJ, Campbell M, Izumiya Y.** 2013. Kaposi's sarcoma-associated herpesvirus (KSHV) latency-associated nuclear antigen regulates the KSHV epigenome by association with the histone demethylase KDM3A. J Virol **87:**6782-6793.
77. **Izumiya Y, Izumiya C, Hsia D, Ellison TJ, Luciw PA, Kung HJ.** 2009. NF-kappaB serves as a cellular sensor of Kaposi's sarcoma-associated herpesvirus latency and negatively regulates K-Rta by antagonizing the RBP-Jkappa coactivator. J Virol **83:**4435-4446.
78. **Izumiya Y, Ellison TJ, Yeh ET, Jung JU, Luciw PA, Kung HJ.** 2005. Kaposi's sarcoma-associated herpesvirus K-bZIP represses gene transcription via SUMO modification. J Virol **79:**9912-9925.
79. **Tischer BK, Smith GA, Osterrieder N.** 2010. En passant mutagenesis: a two step markerless red recombination system. Methods Mol Biol **634:**421-430.
80. **Brulois KF, Chang H, Lee AS, Ensser A, Wong LY, Toth Z, Lee SH, Lee HR, Myoung J, Ganem D, Oh TK, Kim JF, Gao SJ, Jung JU.** 2012. Construction and manipulation of a new Kaposi's sarcoma-associated herpesvirus bacterial artificial chromosome clone. J Virol **86:**9708-9720.
81. **Fakhari FD, Dittmer DP.** 2002. Charting latency transcripts in Kaposi's sarcoma-associated herpesvirus by whole-genome real-time quantitative PCR. J Virol **76:**6213-6223.
82. **Nesvizhskii AI, Keller A, Kolker E, Aebersold R.** 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem **75:**4646-4658.
83. **Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK.** 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun **10:**1523.

Figures

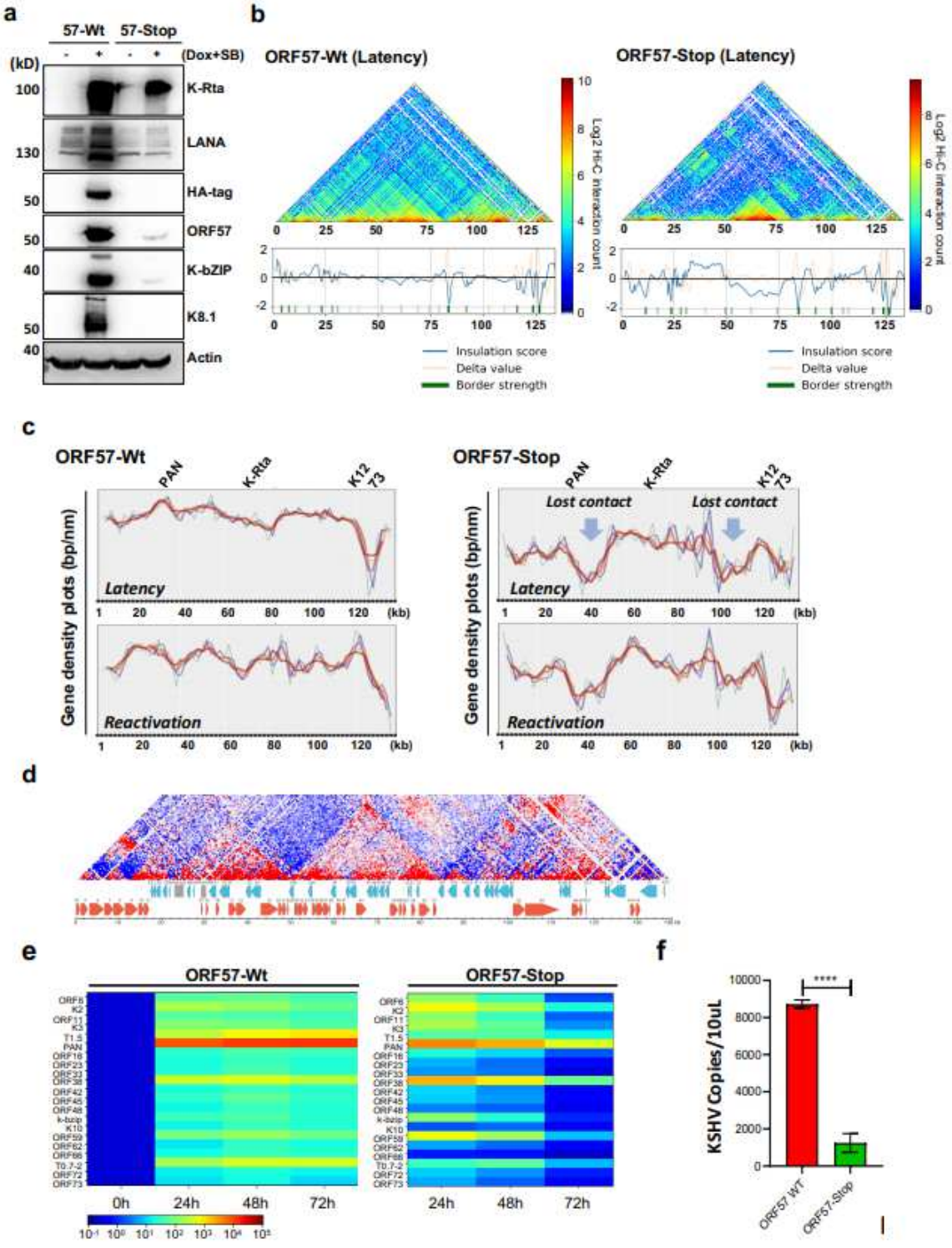

Figure 1

KSHV RNA-binding protein functions in genomic looping. (a) Immunoblotting. A recombinant KSHV harboring stop codons in ORF57 open reading frame was generated (ORF57-Stop KSHV) with BAC16. Knock-out of ORF57 protein tagged with a HA epitope was confirmed and the expression of other viral

proteins was also confirmed with the indicated antibodies. (b) Frequencies of genomic looping. Frequencies of genomic interactions within KSHV genomes in ORF57-Wt (left) or ORF57-Stop (right) KSHV infected cells were plotted by TADbit as heatmap (Bin: 0.5 kb). Insulation score and border strength of TADs in KSHV genomes were also analyzed and plotted in the bottom of panels. (c) Gene-density plots and relative changes in genomic looping by ORF57 knock-out. Gene densities, which indicates frequencies of neighboring to other genes, are shown (Bin: 2 kb). Differences in the genomic loops between ORF57-Wt and ORF57-Stop in reactivated cells were calculated by subtracting from relative abundances of genomic loop in ORF57-Stop and shown as heatmap (bottom panel). The heatmap was plotted with Juicebox. (d) 3C analyses. The PAN RNA genomic region was used as anchor sites to perform 3C with qPCR. KSHV BAC DNA was used as control. (e) RT-qPCR. Expression of selected KSHV genes from different kinetics classes was measured by RT-qPCR and relative abundance was visualized as a heatmap. The amount of transcripts at 0 hr time point was set as the baseline (i.e. 1). (f) Viral replication. KSHV virion production was measured by encapsidated KSHV DNAs in culture supernatant after 96 hr reactivation with K-Rta induction with TPA. **** $P<0.0001$.

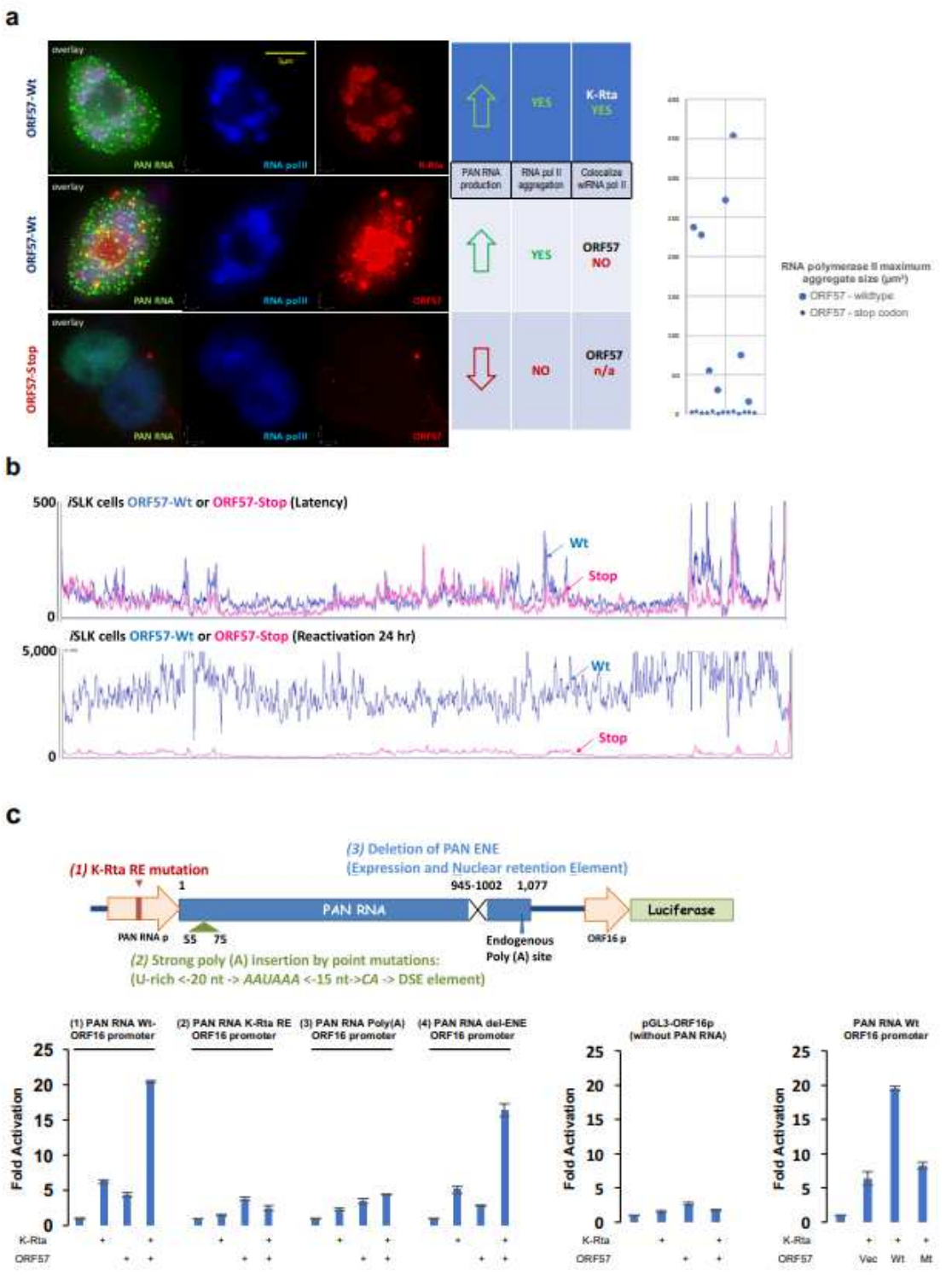

Figure 2

Induction of RNA polymerase II aggregation on KSHV genomes. (a) RNA-FISH and IFA. Recombinant KSHV infected iSLK cells were induced for reactivation by induction of K-Rta expression and TPA treatment for 28 hours. PAN RNA, RNA pol II, and ORF57 protein were visualized by either specific oligo-probes or antibodies (anti-HA tag for ORF57). Volumes of aggregates and intensity of signals were measured with Velocity software. (b) RNA pol II recruitment. CUT&RUN was performed with an anti-RNA

pol II antibody. Occupancy sites were visualized with the Integrative Genomics Viewer (IGV) and screenshots are presented that depict RNA pol II recruitment before (latency; top panel) and 28 hours after reactivation (bottom panel). Blue lines; ORF57-wt cells, Pink line; ORF57-Stop cells. (c) Distal promoter activation by PAN RNA expression. A schematic diagram of the reporter constructs is shown. Fold activation over vector control transfected is provided in bar charts. Total amount of plasmids transfected was normalized by supplementing empty vector and assays were performed with HEK293 cells in triplicated wells. Data are means ± standard deviation of triplicate samples from a single experiment and are representative of three independent experiments (c).

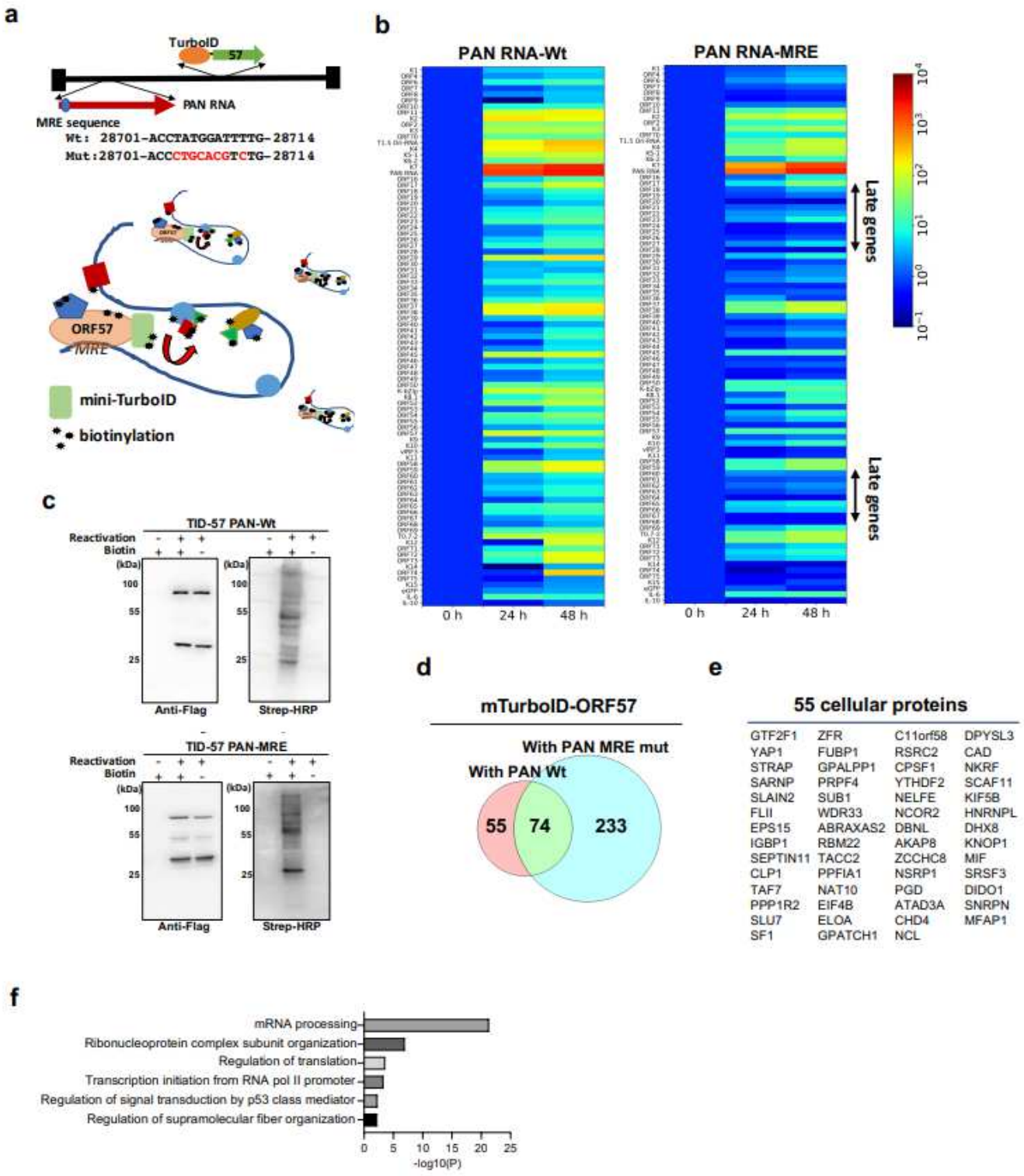

Figure 3

Proximity biotin labeling with mini-TurboID-ORF57 to identify PAN RNA binding proteins. (a) A schematic diagram of recombinant KSHVs. Mini-TurboID is fused to the Nterminus of ORF57 open reading frame and the ORF57 binding site (MRE) was mutated as shown. (b) KSHV RT-qPCR arrays. The genome-wide expression of KSHV genes and lncRNAs was examined by PCR arrays and fold activation over 0 hr time point was plotted as heat maps. (c) Immunoblotting. Recombinant KSHV infected iSLK cells were either

induced reactivation with doxycycline + TPA for 24 hours, protein was probed with either anti-Flag antibody (for miniTurboID-ORF57) or Streptavidin-HRP (for biotinylated proteins). (d) Venn diagram. Venn diagram indicates the number of identified proteins with a P value less than 0.05. Each protein ID was performed in three biological replicates. (e) PAN RNA-mediated interacting proteins. Selected protein names were listed. The entire list of proteins and peptide counts is also presented in Extended Data Table 1. (f) Gene Ontology analyses of PAN RNA-interacting proteins. Non-redundant top 5 biological pathways are presented.

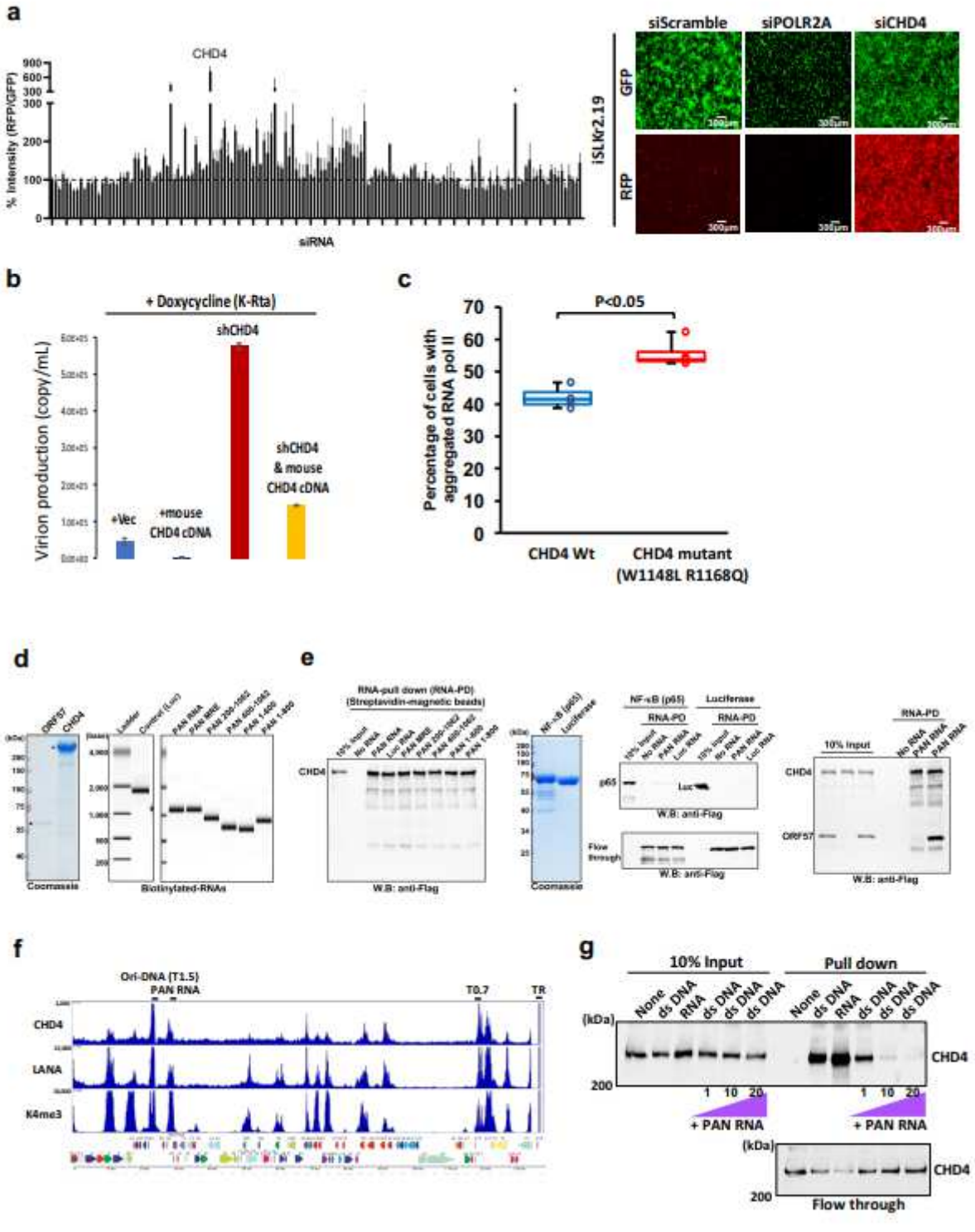

Figure 4

CHD4 is an RNA and DNA binding protein and a strong silencer of KSHV gene expression. (a) siRNA screening with KSHV reactivation as readout. Recombinant KSHV reactivation, which encodes RFP under control of the PAN RNA promoter, was used to screen the effects of KSHV reactivation. RFP signal intensity was measured with Image J, and the GFP signal was used as internal controls. Three randomly selected fields in the middle of each wells were quantified and the average intensity was plotted. Representative pictures were also shown in right panels. siPOLR2A was used as a positive experimental control. (b) CHD4 overexpression and complementation. Encapsidated KSHV genomes in the culture supernatant were measured by qPCR to assess effects of CHD4 expression on KSHV replication. CHD4 was over expressed, knocked-down, or knocked-down then compensated by transfection of mouse CHD4 cDNA or transduction of shRNA. (c) CHD4 ATPase activity is important for prevention of nuclear aggregate formation. Flag tagged mouse CHD4 or mutant was transfected and stained with anti-Flag. Transcriptional factory formation was assessed by staining RNA polymerase II. The percentage of aggregate formation in Flag positive cells was measured and plotted. Three biological replicates and counted 50 cells were counted in each of three biological replicates. (d) Recombinant proteins and synthesized PAN RNAs. Recombinant proteins were expressed with baculoviruses and purified with Flag-agarose beads. Coomassie staining of Flag-ORF57 and Flag-hCHD4 are shown. (e) CHD4 binds RNAs. RNA pull down was performed with indicated biotinylated RNAs and visualized by immunoblotting with anti-Flag antibody. Beads alone was used for background control. ORF57 protein did not compete or enhance CHD4 RNA binding (right panel). (f) CHD4 binding sites on the KSHV genome. CUT&RUN was performed with BC-1 by using indicated antibodies and sequence reads were mapped to KSHV genome. Snapshot of IGV viewer and KSHV genome map were shown. Results of similar experiments with BCBL-1 and BC3 were presented in Expanded Data Figure 4. (g) PAN RNA competes with CHD4 binding on PAN RNA genomic region. Pulldown analyses with biotinylated ssRNA or dsDNA was performed. CHD4 (100 nM) was incubated with 100 nM of biotinylated RNA or 100 nM of biotinylated dsDNA in 40 uL binding buffer. Increased amount of non-biotinylated PAN RNA at 1:1, 1:10, and 1:20 (dsDNA vs. ssRNA) was also incubated and precipitated CHD4 protein in the pulldown was probed with antiFlag antibody. Flow-through and 10% of reaction before pull-down were used as control. An equal amount of dsDNA on beads was also confirmed by qPCR. Data are means ±SEM of triplicate samples from a single experiment and are representative of three independent experiments (a, b).

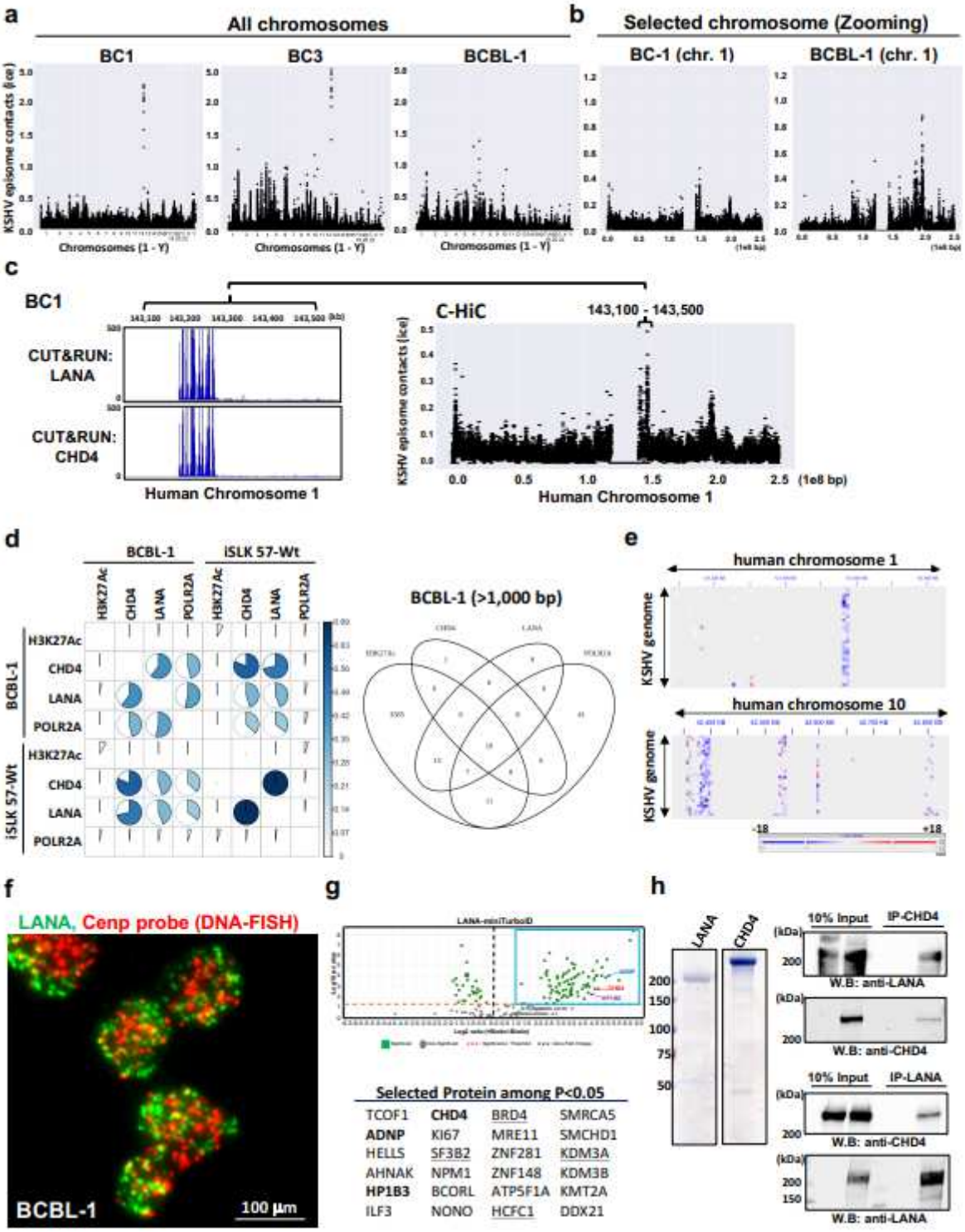

Figure 5

KSHV episomes colocalize with ChAHP complex on host chromatins. (a) KSHV episome docking sites on host cell chromosomes. CHi-C chimeric DNA ligation products composed of sequences derived from the KSHV and human genomes were mapped in three PEL cells. For BCBL-1 cells, three biological replicates were performed with nearly identical results (similarity 0.95). Reads counts of sum of 10 kb Bin were plotted. (b) Zoomed selected chromosome view. Selected chromosome views were also depicted, which

shows enrichment of sequence reads near centromeres. Extended panels were also presented in Extended Data Figure 4. (c) Association of KSHV episomes docking sites with LANA and CHD4 recruitments. CUT&RUN sequence reads with anti-LANA or anti-CHD4 were mapped and visualized with IGV. A snapshot of the IGV view is presented. (d) Jaccard statistics and Venn diagram. CUT&RUN peaks spanning more than 1 kb were isolated and analyzed co-localization on host chromosome with Jaccard statistics. Venn diagram showed the intersection of genomic regions with the length more than 1,000 bp of 4 different proteins and the Jaccard statistic was used to compare similarity of positions of those genes. (e) A detachment of KSHV episomes by reactivation. Relative C-HiC reads were visualized by Juicebox as similar to Fig. 1c. (f) KSHV episomes locates near the centromere. DNA-FISH was used to visualize location of centromere and KSHV episomes were located with anti-LANA antibody with IFA. Two signals were frequently localized next each other. The 3D view was presented in Extended Data Figure 8. (g) mini-TurboID LANA. Recombinant KSHV, which harbors mini-TurboID fused LANA was generated and infected in iSLK cells. Protein localizing in the proximity of LANA were identified with a mass spectrometry. Selected proteins with p value less than 0.05 were listed. Each protein ID was performed in three biological replicates. T-test was used for calculating the p value. Bold case; ChAHP components. Underlined; previously identified proteins as LANA interacting proteins. (h) LANA interacts with CHD4 in vitro. Purified LANA and CHD4 were prepared from recombinant baculovirus infected Sf9 cells. Coomassie staining is shown. The equal molar amount (100 nM) of two proteins were incubated in binding buffer and immunoprecipitated with the specific antibodies of one another. The 10% of reaction before immunoprecipitation was shown as controls.

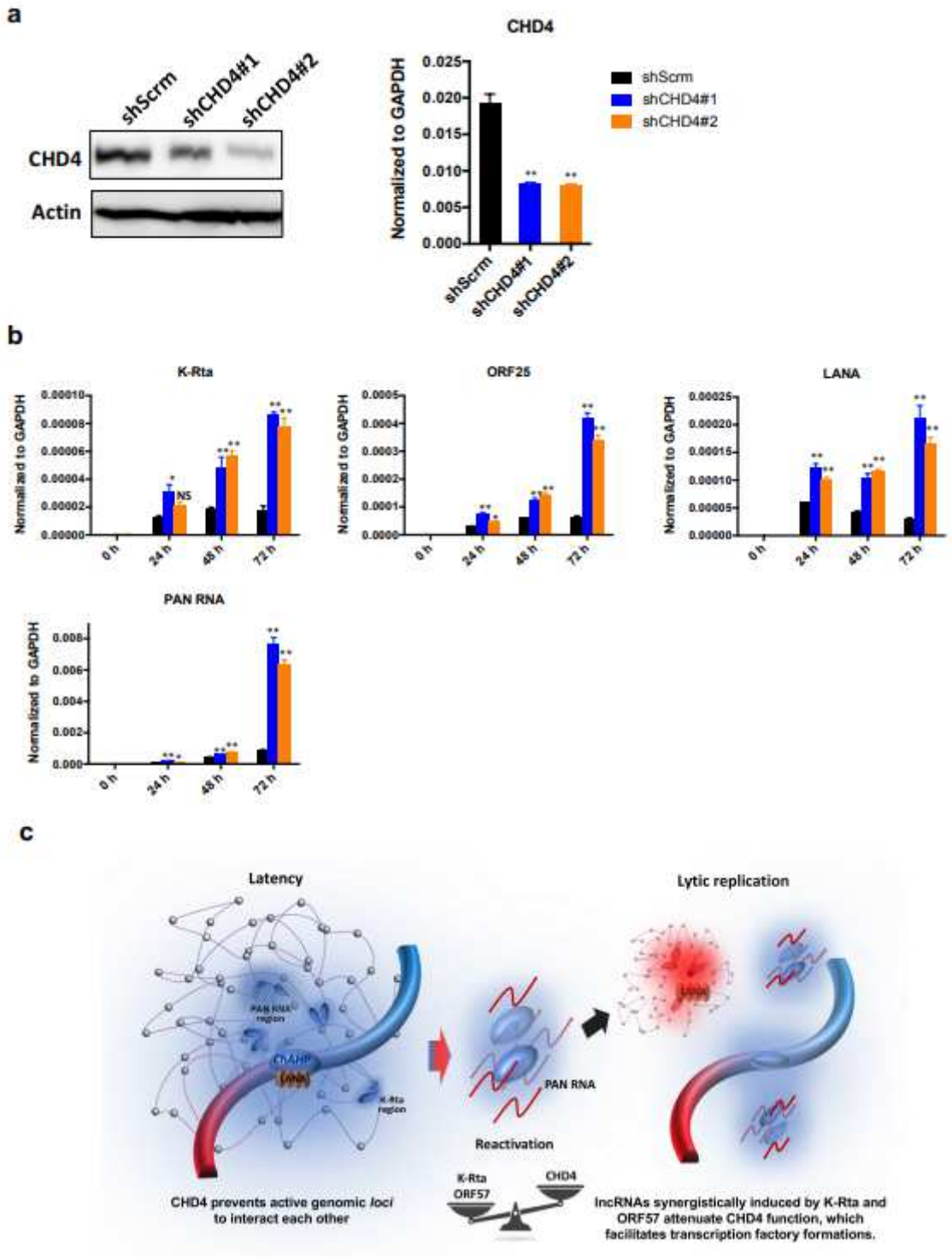

Figure 6

CHD4 is important for establishing latency. (a) CHD4 knockdown and KSHV lytic gene expression during de novo infection. Immunoblot analysis of CHD4 protein and RT- qPCR analysis of CHD4 mRNA expression in scramble shRNA- (shScrm) or shCHD4- transduced stable 293T cell lines are shown. 293T cells stably transduced with shScrm or shCHD4 were infected with rKSHV.219 for 24 h. Total RNA was harvested at the indicated time points. K-Rta and ORF25 expression was measured by RT-qPCR and

normalized to GAPDH. * P<0.05, ** P<0.01 (comparing with shScramble control). (b) KSHV latency model. KSHV episomes tether at borders of host TADs with ChAHP complex. Stimulation of K-Rta, in part by decreased CHD4 expression, triggered robust PAN RNA expression, which compete CHD4 DNA binding to facilitate KSHV reactivation. Leaving CHD4 enriched nuclear environment by detaching from host chromatin further enhance KSHV reactivation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 2932150supp173666qm0g0y.pdf
- 2932150dataset172636qlgmb4.xls
- 2932150dataset172637qlgmb4.xls
- 2932150dataset172638qlgmb4.xls