KSHV Topologically Associating Domains in Latent and Reactivated Viral Chromatin

Mel Campbell[1,*], Chanikarn Chantarasrivong[2,*], Yuichi Yanagihashi[3], Tomoki Inagaki[1], Ryan R. Davis[4], Kazushi Nakano[2], Ashish Kumar[1],Clifford G. Tepper[5, 6], & Yoshihiro Izumiya[1, 5, 6, #]

[1] Department of Dermatology, School of Medicine, University of California Davis (UC Davis) Sacramento, CA 95817 USA,
[2] Lifescience Division, Lifematics, Osaka, 541-0046 Japan
[3] Biological Data-Science Division, Lifematics, Osaka, 541-0046 Japan
[4] Department of Pathology and Laboratory Medicine, School of Medicine, UC Davis, Sacramento, CA 95817 USA,
[5] Department of Biochemistry and Molecular Medicine, School of Medicine, UC Davis, Sacramento, CA 95817 USA,
[6] UC Davis Comprehensive Cancer Center, Sacramento, CA 95817 USA

* Equal contribution.

#Correspondence and request for materials should be addressed to Y.I.

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

Word count:

Abstract: 200

36 **Abstract**

37 Eukaryotic genomes are structurally organized via the formation of multiple loops that

38 create gene expression regulatory units called topologically associating domains

39 (TADs). Here we revealed the KSHV TAD structure at 500 base pair resolution and

40 constructed a 3D KSHV genomic structural model with 2kb binning. The latent KSHV

41 genome formed very similar genomic architectures in three different naturally infected

42 PEL cell lines and in an experimentally infected epithelial cell line. The majority of the

43 TAD boundaries were occupied by CTCF and SMC1, and the KSHV transactivator was

44 recruited to these sites during reactivation. Triggering KSHV gene expression

45 decreased pre-wired genomic loops within the regulatory unit, while contacts extending

46 outside of regulatory borders increased, leading to formation of a larger regulatory unit

47 with a shift from repressive to active compartments (B to A). The 3D genomic structural

48 model proposes that the immediate-early promoter region is localized on the periphery

49 of the 3D viral genome during latency, while highly inducible non-coding RNA regions

50 moved toward the inner space of the structure, resembling the configuration of a "bird

51 cage" during reactivation. The compartment-like properties of viral episomal chromatin

52 structure and its reorganization during the transition from latency may help coordinate

53 viral gene transcription.

54

55 **Importance**

56 The 3D architecture of chromatin allows for efficient arrangement, expression, and

57 replication of genetic material. The genomes of all organisms studied to date have

58 been found to be organized through some form of tiered domain structures. However,

2

59 the architectural framework of the genomes of large double-stranded DNA viruses such

60 as the herpesvirus family has not been reported. Prior studies with Kaposi's sarcoma-

61 associated herpesvirus (KSHV) have indicated that the viral chromatin shares many

62 biological properties exhibited by the host cell genome, essentially behaving as a mini

63 human chromosome. Thus, we hypothesized that the KSHV genome may be organized

64 in a similar manner. In this report, we describe the domain structure of the latent and

65 lytic KSHV genome at 500 base pair resolution and present a 3D genomic structural

66 model for KSHV under each condition. These results add new insights into the complex

67 regulation of the viral lifecycle.

68

Introduction

Kaposi's sarcoma-associated herpesvirus (KSHV), also known as human herpesvirus-8 (HHV8), is a member of the gammaherpesvirus family of double-stranded DNA viruses. The virus is strongly associated with Kaposi's sarcoma (KS), an endothelial cell-derived tumor, and two rare lymphoproliferative disorders, multicentric Castleman's disease (MCD), and primary effusion lymphoma (PEL). KSHV exhibits a broad host range and can infect a variety of cell types *in vitro*, including B lymphocytic cells, renal-derived cells, and human gingival epithelial cells (1-3). KSHV DNA exists in the virion as a linear duplex of ~140 kb of coding sequence which encodes approximately 80 genes or open-reading frames (ORFs) that are expressed in a highly-coordinated manner during latency or lytic replication (4, 5). The viral coding sequences are flanked on either side by tandem terminal repeats of highly GC-rich noncoding sequences giving rise to a genome of ~160-170 kb (6, 7). Following infection, the chromatin-free viral genome circularizes, is rapidly chromatinized (8) and maintained and replicated in the nucleus of the host cell as monomeric episomes. A crucial element of the KSHV lifecycle is the reactivation of the virus from a latent (dormant) state into the lytic replicative cycle in which viral genomic DNA is replicated, viral particles are produced, and then virions are released along with host cell lysis.

A great deal of effort has focused on defining the epigenetic mechanisms that regulate latency, reactivation, and commandeering of the host cell by KSHV (for reviews, see (9-11). Although a single viral protein, K-Rta (ORF50), triggers the onset of lytic replication by the transcriptional activation of lytic genes, the exact details of how K-Rta functions as this master regulator have been enigmatic. We have previously shown

92   that interplay between KSHV-encoded transcription factors and host cell-encoded

93   epigenetic regulators (JMJD2A demethylase, EZH2 methyltransferase) are key

94   mechanisms in controlling KSHV reactivation (8, 12). In addition, using Capture Hi-C

95   (chromosome conformation capture analyses with deep sequencing) analyses of the

96   KSHV genome, we demonstrated that K-Rta not only activates individual lytic genes by

97   binding to specific regulatory elements along the KSHV genome, but that it also

98   performs a higher-order coordination of the process by mediating three-dimensional

99   (3D) conformational changes in the architecture of the KSHV chromosome (13). In

100   essence, "chromatin loops" are formed via K-Rta promoting contact between distant

101   genomic regions containing key regulatory sequences.

102   Advances in chromosome conformation capture (3C)-based studies have

103   uncovered the existence of spatially insulated genomic regions that are now considered

104   the invariant building blocks of chromosomes (14-17). When initially reported in 2012,

105   these ~100 kb-1 Mb regions were referred to as topological domains (14) or

106   topologically associating domains (TADs) (15). These domains are defined by the

107   preferential interaction of loci located within a given TAD and a relative (~2-fold)

108   depletion of interactions between loci located in different TADs (16). In the era of 3C-

109   based methods, TADs were originally described in mammalian cells (14, 15) using high-

110   throughput chromosome conformation capture (Hi-C) (18) or chromosome conformation

111   capture carbon-copy (5C) (19), and in *Drosophila* (17) chromosomes using Hi-C.

112   However, TADs have also been found in Zebrafish (20), *Caenorhabditis elegans* (21),

113   *Saccharomyces cerevisiae* (22) and *Schizosaccharomyces pombe* (23). In addition,

114   analysis of nucleoids of *Caulobacter crescentus* (24) and *Bacillus subtilis* (25, 26) have

115   described genomic spatial domains termed Chromosomal Interaction Domains (CIDs)

116   which mimic TADs in terms of preferential interaction properties. Together, these results

117   suggest that TAD-like domains, although differing in size among the genomes of the

118   various species analyzed (27), may be a universal feature of both prokaryotic and

119   eukaryotic genomes.

120   The existence of TAD-like structures within the genomes of large double-

121   stranded DNA viruses such as herpesvirus family has not been reported. For KSHV

122   there are numerous reports concerning the role of the structural maintenance of

123   chromosomes (SMC) cohesin complex and CCCTC-binding factor (CTCF) in control of

124   viral latency and reactivation (28-36). As both cohesin and CTCF have been linked to

125   chromatin looping formation in mammalian cells (14, 37-39), this suggests the plausible

126   existence of TAD-like structures within KSHV genomes. Moreover, previous 3C (33)

127   and capture Hi-C (13) analyses have documented the presence of chromatin looping

128   within the latent and lytic KSHV genomic regions. Together with KSHV epigenomic

129   mapping data (8, 12, 40-43), this suggests that the chromatinized KSHV genome

130   shares many biological properties exhibited by the host genome, behaving as a mini

131   human chromosome and, as such, may be organized in a TAD-like manner. In this

132   report, we describe the TAD-like structures of the latent and lytic KSHV genome at 500

133   base pair (bp) resolution and present a 3D genomic structure model for the KSHV under

134   each condition.

135

136

137   **Materials and Methods**

138   **Capture Hi-C**: KSHV Capture Hi-C (CHi-C) analysis was performed using a robust *in*

139   *situ* CHi-C protocol with kitted reagents from Arima Genomics (San Diego, CA) based

140   on well-accepted methods described for *in situ* Hi-C (38, 44), CHI-C (45-47), and as

141   described previously (13). Briefly, cells were crosslinked with 2% formaldehyde, lysed,

142   and genomic DNA digested with a cocktail of 4-cutter restriction endonucleases by

143   incubation for 30 minutes at 37°C. The 5'-overhangs were then filled in and "marked"

144   with biotin by incorporating biotinylated dATP (biotin-14-dATP) with Klenow fragment of

145   DNA polymerase I (incubation for 45 minutes at 25°C). Proximity intramolecular ligation

146   of the blunt-ended fragments was then performed with T4 DNA ligase (incubation for 15

147   minutes at 25°C). The formaldehyde crosslinks were reversed, and the ligated, chimeric

148   DNA products were purified with Agencourt AMPure XP paramagnetic beads (Beckman

149   Coulter). The DNA was then sheared to an average length of 400 bp using a Covaris

150   E220 Focused-ultrasonicator (Covaris, Inc.) and fragments size-selected with AMPure

151   XP beads to achieve a size distribution of 200-600 bp. The biotin-marked ligation

152   products were then enriched by affinity capture with streptavidin magnetic beads

153   (DynaBeads MyOne Streptavidin C1; Invitrogen, Thermo Fisher Scientific).

154   Subsequently, libraries were prepared from the bound ligation products with the Kapa

155   HyperPrep Kit with Library Amplification Module (Roche) using an on-bead modification

156   to the standard protocol for end repair, dA-tailing, and ligation of Illumina TruSeq

157   sequencing adaptors.

158   The KSHV CHi-C library was then prepared by target enrichment of the libraries

159   for KSHV genomic content by solution hybridization with a custom-designed KSHV

7

160   genomic capture probe library (xGen Lockdown Probes; Integrated DNA Technologies,

161   Inc., Coralville, IA) as previously described (13). Briefly, libraries (500 ng) were

162   hybridized with the KSHV genomic capture probe pool (3 pmol) in a mixture containing

163   xGen 1X Hybridization Buffer, Cot-1 (5 µg), and xGen Universal Blocking Oligos for 4

164   hours at 65°C. The hybridized targets were then captured with streptavidin beads

165   (DynaBeads MyOne Streptavidin C1; Thermo Fisher Scientific) by incubation for 45

166   minutes at 65°C. Unbound DNA was removed by a series of high-stringency (65°C) and

167   low-stringency (room temperature) washes. The KSHV genome-enriched CHi-C library

168   DNA was eluted and PCR enrichment (12 cycles) performed with high-fidelity KAPA

169   HiFi HotStart DNA Polymerase (Kapa Biosystems, Inc., Wilmington, MA). Libraries were

170   multiplex sequencing (2 x 150 bp, paired-end, ~30 million mapped reads/mate pairs per

171   sample) on an Illumina Hiseq 4000 sequencing system (Supplement Figure S1).

172

173   **Hi-C data pre-processing**: Sequence alignment and quality check of reads were

174   performed using HiC-Pro 2.11.1 pipeline (48); each read-end was aligned to the human

175   GRCh37 (hg19) and KSHV (NC_009333.1) reference genomes. Quality reports showed

176   percentage duplicates at less than 15% while >95% di-tags were valid and aligned to

177   restriction fragments (S-Figure 1). The duplicates were removed and reads mapping only

178   to the KSHV genome were extracted by using Pysam 0.14.1.

179

180   **Genomic domain analysis**: The KSHV genome was analyzed and visualized for

181   frequencies of chimeric genomic reads, and finally modeled for 3D Structure using

182   TADbit (49). The KSHV mapped reads were filtered for only uniquely mapped reads

8

183   pairs based on the intersection of each read-end. The reads that were self-ligations,

184   dangling-end, error, extra dangling-end, too short, too large, duplicated, and random

185   breaks were filtered out to provide valid reads. The valid reads were stored as matrices

186   and binned with resolution of 500 bp (2 kb for 3D modelling); the bins with more than

187   1000 counts and at least 75% of cells with no-zero counts were used in the next steps.

188   Iterative Correction and Eigenvector decomposition (ICE) normalization was used to

189   treat the data with iteration=100. Bins were identified as in compartment A or B by

190   calculating eigenvectors. TADs/TAD borders were called, and insulation score/border

191   strength were calculated using the TADbit computational framework (49), and visualized

192   as contact maps and aligning plots.

193

194   **KSHV 3D chromosome structure modelling:** Matrix Modeling Potential (MMP) score

195   (50), ranging from 0 to 1, was used to identify if the interaction matrices from the

196   previous step (i.e., iterative modeling with TADbit), have potential for modeling. The

197   higher the MMP scores are, the more likely the 3D model will be accurate. In this study,

198   MMP scores of latency and reactivation models were 0.9541 and 0.9023, respectively.

199   Monte Carlo optimization method was used to optimized parameters from both models.

200   The parameters included maximal distance associated between two interacting

201   particles, particles attraction, particles repulsion, and contact distance of particles. Sets

202   of models were produced from the possible combinations of those parameters. Contact

203   maps for each set of models were built, then compared with the Hi-C interaction by

204   averages of Spearman correlation coefficients. The models with higher correlation

205   coefficients represented the original data and were chosen to be visualized and plotted

206   by DNA density. The correlations of the latency model and reactivation models were

207   0.8706 and 0.8782, respectively. Finally, three-dimensional visualization of KSHV

208   molecular modeling were performed with the UCSF Chimera package (51).

209

210   **Cleavage Under Targets and Release Using Nuclease (CUT&RUN).** CUT&RUN (52)

211   was performed essentially by following the online protocol established by Dr. Henikoff's

212   lab with several modifications. Cells were washed with PBS and wash buffer [(20 mM

213   HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine (Sigma, S2626) and

214   proteinase inhibitor (Roche)]. After removing the wash buffer, cells were captured on

215   magnetic ConA beads (Polysciences, PA, USA) in the presence of $CaCl_2$. Beads/cells

216   complexes were washed 3 times with digitonin wash buffer (0.02% digitonin, 20 mM

217   HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine and 1x proteinase inhibitor),

218   aliquoted, and incubated with specific antibodies in 250 μL volume. The antibodies used

219   in this study were: rabbit polyclonal anti-SMC1 (Cell Signaling, #4802S; 1:100), mouse

220   anti-RNA polymerase II (Millipore, clone CTD4H8; 1:100), and rabbit monoclonal anti-

221   CTCF (Cell Signaling, clone D31H2, #3418S; 1:100). After incubation, unbound

222   antibody was removed by washing with digitonin wash buffer 3 times. Beads were then

223   incubated with recombinant pAG-MNase, which was purified from *E.coli* (S-Figure 1), in

224   250 μl digitonin wash buffer at 1.0 μg/mL final concentration for one hour at 4°C with

225   rotation. Unbound pAG-MNase was removed by washing with digitonin wash buffer 3

226   times. Pre-chilled 2 mM $CaCl_2$ containing digitonin wash buffer (200 μL) was added to

227   beads and incubated on ice for 30 min. The pAG-MNase digestion was halted by the

228   addition of 200 μl 2× STOP solution (340 mM NaCl, 20 mM EDTA, 4 mM EGTA,

229   50 μg/ml RNase A, 50 μg/ml glycogen). The beads were incubated with shaking at 37°C

230   for 10 min in a tube shaker at 500 rpm to release digested DNA fragments from the

231   insoluble nuclear chromatin. The supernatant was collected after centrifugation (16,000

232   x g for 5 min at 4°C) and placed on a magnetic stand. DNA was extracted using the

233   NucleoSpin kit (Takara Bio, Kusatsu, Shiga, Japan). Sequencing libraries were then

234   prepared from 3 ng of CUT&RUN DNA with the Kapa HyperPrep Kit (Roche) according

235   to the manufacturer's standard protocol. Libraries were multiplex sequenced (2 x 150bp,

236   paired-end) on an Illumina HiSeq 4000 sequencing system to yield ~15 million mapped

237   reads per sample. When necessary, *E. coli* genomic DNAs carried over in purified pAG-

238   MNase was used to normalize sequence reads as described previously (52).

239

240   **Genomic data analysis**.

241   FASTQ files for the capture Hi-C experiments were processed through the HiCUP

242   (v0.7.4) pipeline (53) using the KSHV (NC_009333.1) genome. Valid interaction

243   products called by HiCUP were converted into Juicebox (54) input format (.hic file),

244   which stores the normalized and un-normalized contact matrices as a highly

245   compressed binary file, by using a series a scripts provided by HiCUP (hicup2homer)

246   and HOMER (makeTagDirectory and tagDir2hicFile) (55). Juicebox was utilized to

247   facilitate adjustments of resolution and normalization, intensity scaling, zooming, and

248   addition of annotation tracks.

249   CUT&RUN sequence reads were aligned to the human hg38 reference genome

250   assembly and reference KSHV genome sequence with Bowtie2 v2.3.5.1 (56) and/or

251   HISAT2 v2.1.0 (57). MACS2 (Model-based Analysis of ChIP-Seq) v2.1.1.1.20160309

11

252   was used for detecting peaks (58) following the guidelines in the developer's manual.

253   The following publicly available data sets from previous studies (59, 60) were also used

254   to overlay binding sites or histone modification sites; KSHV K-Rta ChIP-seq

255   (GSE123897), H3K27Ac, H3K4me3, and H3K27me3 (GSE163695). Peaks and read

256   alignments were visualized using the Integrated Genome Browser (IGB) (61). Heatmaps

257   and average profile plots were drawn from bed files created by MACS and bam files

258   using the R package, ngsplot v2.63 (62).

259

260   **Cell Culture, bacmid transfection, and selection**

261   TREx-F3-H3-K-Rta BCBL-1 cells (TREx-BCBL-1) cells were maintained and reactivated

262   as previously described (13). BC-1 and BC-3 cells were maintained in RPMI 1640

263   supplemented with 15% FBS, 2mM glutamine, and 1% penicillin-streptomycin. BC-1

264   and BC-3 cell lines were obtained from the ATCC (Manassas, Va). iSLK cells that were

265   infected with BAC16 HA-ORF57 Wt virus, were maintained in DMEM supplemented

266   with 10% FBS, 1% penicillin-streptomycin-L-glutamine solution, 1 mg/ml hygromycin B,

267   and 250 µg/ml G418. iSLK cells were obtained from Dr. Don Ganem (Novartis Institute

268   for Biomedical Research) and were maintained in DMEM supplemented with 10% FBS,

269   1% penicillin-streptomycin-L-glutamine solution and 400 µg/ml hygromycin and 250

270   µg/ml G418 and 10µg/ml puromycin.

271

Results

KSHV topologically associating domains

In our previous study, we demonstrated that the KSHV genome forms organized DNA loops in infected cell nuclei and the frequency of these loops increased during reactivation, especially near K-Rta binding sites (13, 63). In order to understand the molecular details of loop formation, we have performed additional Capture Hi-C experiments to augment the results from our earlier publication. Previously, our experimental workflow utilized an in-house 3C procedure based on Gavrilov et al. (64) with a single restriction enzyme digestion prior to library capture with KSHV lockdown probes (13). In the current study, a commercially available Hi-C kit (Arima Genomics) was used in conjunction with our probe library. In this procedure, the chromatin digestion step is performed *in situ* and employs a restriction enzyme cocktail coupled with fill-in reactions to incorporate a biotinylated deoxynucleotide to mark the digested ends which facilitates the enrichment of the DNA ligation products (i.e., consisting of the spatially proximal digested ends) prior to the KSHV lockdown probe capture step. The restriction enzyme digestion generated fragment sizes with an average length of 190 bp +/- 210 bp. The largest fragment is 2,225 bp, which is located near the LANA coding region, while the shortest fragment is 4 bp. The distribution of fragment sizes is shown in S-Figure 2B. A total of 3,565,321 valid read pairs were uniquely mapped on the KSHV genome in non-reactivated BCBL-1 cells, while 5,967,349 valid read pairs were mapped in the reactivated sample. Using two capture/enrichment steps, we further increased the number of sequence reads corresponding to valid Hi-C di-tags by approximately 1,000-fold over our previous studies and prepared KSHV genomic

13

295   interaction maps at 500 bp resolution (Figure 1, S-Figure 2). Because topologically

296   associating domains often refer to sizes of 100 kb to 1 Mb genomic domains (14), we

297   decided to use the term of transcription regulatory domain (TRD).

298   The KSHV episomes present in three naturally infected PEL cell lines (BCBL-1,

299   BC1 and BC3) all exhibited very similar TRD structures (Figure 1). In addition, we also

300   determined KSHV TRD structure in experimentally-infected *i*SLK/Bac16 cells. The

301   results demonstrated very similar TRD structures to those of the naturally-infected cells.

302   Our TRD map also agreed with previous 3C studies that positioned the LANA promoter

303   region forming genomic loops with K12 loci at higher frequencies (31). By overlaying

304   previously determined histone modification and transcription factors binding sites, we

305   next examined the association with TRD formation (Figure 2). Although the genomic

306   domains are not defined by chromatin state, sites of histone modification aligned very

307   well with TRD formation and were organized with genomic domains with active histone

308   marks (H3K4me3, H3K27Ac). In fact, the genomic regions with active histone marks

309   seem to be physically neighboring each other in the 3D structure, which is seen by

310   higher frequencies of genomic ligations than the theoretically-expected ligated

311   sequence reads (shown in red color, where two H3K4me3 marked genomic regions

312   were ligated more frequently than theoretically-expected frequency), and these features

313   were common in all cell lines tested. The results also demonstrated that TRDs could

314   also be distinctly separated based on the reactivation kinetics class of KSHV gene

315   expression. For instance, late gene clusters were localized in distinct TRDs (within

316   separated triangles). Consistent with previous reports, CTCF localizes to the junctions

317   of the majority of TRDs, and CTCF indeed colocalized with SMC1 on the KSHV

318   genomes (Figure 2B, C). In addition, superimposing K-Rta recruitment sites that were

319   reported previously (59) further showed that K-Rta recruitment sites (Figure 2C black

320   line) localized adjacent to CTCF/SMC1 binding sites (Figure 2C, red and blue lines),

321   and the genomic regions harbored poised/stalled RNAPII (Figure 2C shadow). The

322   results indicate that the KSHV episome is structured to allow the K-Rta complex to

323   access to other genomic regions through genomic looping, which may serve to

324   maximize K-Rta transactivation function. The NGS plot further confirmed that K-Rta

325   binding sites are closely localized to CTCF and RNAPII binding sites (Figure 2D).

326

327   **Regulation of TAD structure during reactivation.**

328   Having defined the KSHV genomic TRD structure in the condition that the majority of

329   episomes were in during the latent state, we next examined changes in TRDs induced

330   during the early stages of reactivation. To avoid complications with ongoing DNA

331   replication, we examined a time point (i.e., 24 hrs post-reactivation) at which there was

332   an only minimal increase in viral DNA copies, as measured by qPCR (60). Capture Hi-C

333   analyses were performed with TREx-BCBL-1 cells at 0 and 24 hours post-reactivation

334   using an induction protocol consisting of a combination of TPA and doxycycline. The

335   experiments were performed in triplicate. Using the TAD caller program TADbit for

336   visualization of hierarchical genomic domains, the formation of larger TRDs during

337   reactivation was observed when compared with TRD structures observed during

338   latency, and presumably proceeded through the fusion of neighboring TRDs via

339   assembly of active transcription sites, [Figure 3A (a, b), (63)]. The results also showed

340   that pre-existing genomic loops in latent cells (Figure 1 in red) were largely disrupted

15

341   during reactivation, as indicated by a marked decrease in the frequency of their

342   interactions within individual TRDs (Figure 3B, visualized as blue color). While pre-

343   existing TRDs were disrupted, genomic contacts that spanned outside of the previously

344   existing TRD borders were increased (Figure 3B, visualized as red color). Reactivation

345   was also characterized by new interactions of the K-Rta promoter region formed with

346   the late gene clusters (ORF63-68), while pre-formed, latency-associated contacts with

347   active compartments, such as E gene clusters (ORF6-ORF11) were decreased (Figure

348   3B).

349   Next, insulation scores were calculated to reveal changes in TRDs before and

350   after reactivation. A lower score suggests a higher insulation effect, which is indicative

351   of the position of the boundary. The border strength at the majority of genomic regions

352   was decreased in reactivating samples (Figure 3C), suggesting more freely interaction

353   with other genomic domains when viral lytic gene expression was triggered. As

354   expected, compartment analyses showed shifts from repressive to active compartments

355   (B to A) during reactivation [Figure 3D]. Interestingly, the analyses also showed that E

356   gene clusters encompassing sequences from K1 to PAN RNA coding regions (genomic

357   locations bin 1-60 [1-30 kb region]) showed an active compartment structure in latent

358   chromatin. This result is somewhat unexpected since these genomic loci largely encode

359   lytic genes. While LANA, two Ori-Lyts regions, K10.5/11, and the PAN RNA coding

360   region did not change a specific compartment structure between that observed during

361   latency and after triggering reactivation, these genomic regions are narrowly insulated

362   with higher CTCF/SMC1 peaks (Figure 2B).

363

364   **KSHV 3D genomic structure, a bird cage model.**

365   With 3,565,321 (latent) and 5,967,349 (lytic) valid Hi-C di-tags (i.e., genomic

366   contacts) covering the ~140 kb KSHV genome, we next constructed a theoretical 3D

367   KSHV genomic structural model utilizing TADbit, which was then visualized with the

368   UCSF Chimera program (51). The model proposes that the immediate-early promoter

369   region (e.g., K-Rta) is localized on the periphery of the 3D viral genome during latency,

370   which is likely to be more accessible to the nuclear environment (Figure 4,

371   Supplemental material movie); while the highly inducible long non-coding RNA region

372   (i.e., PAN RNA) moved toward the inner space of the structure in reactivating episomes.

373   During the transition from latency to the lytic cycle, the overall 3D genomic structures

374   were squeezed into a doughnut-like configuration from a spherical shape. This

375   arrangement positions the transcriptionally active regions to be more closely

376   neighboring to other genomic loci in three-dimensional space, and their conformation

377   change approximates a "bird cage" (Figure 4). These changes were also consistent with

378   disruption of pre-existing TADs and the increase in genomic loops extending outside of

379   TAD domains as described in Figure 3A. Taken together, we propose that the KSHV 3D

380   genomic structure is designed to maximize the effects of K-Rta complex recruitment and

381   may also facilitate re-utilization of active RNAPII complexes.

382

383   **Discussion**

384   In this report, we have empirically mapped the genomic contacts within the KSHV

385   genome during latency and at an early stage of reactivation. This mapping, combined

386   with bioinformatic tools, has allowed us to uncover TAD-like structures of the KSHV

17

387   genome during each condition. Using a workflow consisting of HiC-Pro and TADbit, we

388   have constructed a structural model which allows visualization of the overall

389   conformational changes that occur in the viral genome during this time period. With the

390   analyses, similar contact maps were obtained among three PEL cell lines and one

391   laboratory infected cell line (*i*SLK BAC16-Wt). This overall similarity suggests that the

392   prominent viral genome contacts established during latency appear to be conserved

393   among different KSHV infected cell lines and are stable properties within the infected

394   cell population.

395   Previous studies have implicated CTCF sites as being important for several

396   KSHV genomic loops that involve the latency control region and either the K12 or

397   ORF50 locus (31). Inspection of the contact heat maps aligned well with the

398   CTCF/SMC1 recruitment sites at the corners of TRDs, supporting the notion that

399   CTCF/SMC1 plays a major role in defining TRD domains. However, the smaller TRDs

400   also show a lack of CTCF/SMC1 (Figure 2), possibly reflecting that enriched contacts

401   are not readily visible due to the transient nature of extrusion (65) or that other

402   architectural protein factors could also play a role in structuring latent KSHV genomes.

403   In this respect, KSHV latent genome structures resemble "ordinary domains" as

404   previously defined by Rao et al. (38) to classify contact domains not flanked by CTCF

405   binding sites and formed without involvement of CTCF/cohesin-mediated loop extrusion.

406   As those smaller TADs are separated by transcription units (e.g., gene clusters with

407   genes transcribed in the same direction and having similar kinetics), we speculate that

408   RNA-binding proteins may also play a role in viral genomic structure (66). Although we

409   have highlighted contacts between the ORF50 and K12 loci [(13), Figure 4] and a

410 potential lncRNA from this region, it should be noted that it is uncertain which gene

411 products originating from promoters in this region that might be influenced by genomic

412 looping. This region harbors an origin of lytic replication and the lytic ALT lncRNA

413 running antisense (67). Additional possible candidates are mRNAs encoding the

414 Kaposin locus ORFs as well as a lytic miR-K12-10/12 pri-miRNA (4, 68).

415 During the shift from latency to lytic replication viral genomes exhibited a loss of

416 local proximal intra-TRD contacts with the widespread onset of inter-TRD contacts (Fig.

417 3A(a)). Evidence exists for both dynamic (69-71) and stable (71-74) enhancer-promoter

418 interactions during cellular responses to various stimuli or conditions. In different cell

419 lines or across species TADs are generally conserved at the megabase scale (14, 39).

420 However, when probed at higher levels of resolution, TADs may merge or be disrupted

421 during changes in gene expression associated with a variety of cellular events such as

422 development (75), lineage commitment (37), reprogramming (76), differentiation (77,

423 78), senescence (79) and possibly heat shock (80, 81). Together, the current study and

424 our previous report (13), as well as findings from the Lieberman group (31) confirm the

425 existence of both pre-wired and induced genomic loops and the dynamic nature of

426 KSHV genomic contacts during the viral life cycle.

427 Insulation scores revealed the strength of borders on the KSHV genome. CTCF,

428 SMC1, and K-Rta frequently bound at the borders, and these sites were decorated by

429 H3K4me3 or H3K27Ac modifications. This observation suggested that K-Rta targets the

430 boundaries established by CTCF and SMC1 with open chromatin structures. By

431 targeting the boundaries, it is likely that K-Rta can recruit K-Rta-bound chromatin

432 remodeling factors (82) to the multiple pre-assembled genomic regions for synchronous

433   activation. In turn, Mediator complex, which is also a part of the K-Rta activation

434   complex (82, 83), may help to establish active genomic hubs for future reactivation.

435   The border strength of most viral genomic regions decreased during lytic

436   reactivation. This decrease suggests for greater accessibility of a given TRD to interact

437   with other genomic elements, which is accompanied by induction of viral lytic gene

438   transcription. However, the latency-associated loci remain relatively unchanged (Figure

439   3C). The latency genomic region is, in fact, localized in a small valley, which is well-

440   protected by two strong peaks of CTCF and SMC1 in the KSHV genome (Figure 1B;

441   genome coordinates 124-130 kbp). The result indicated that the majority of episomes in

442   an infected cell within the given cell populations have similarly maintained CTCF and

443   SMC1 sites. The result also suggests that these highly-utilized CTCF/SMC1 binding

444   sites are likely to be essential for maintaining KSHV episomes in infected cells.

445   The dynamic nature of the bird cage model is consistent with results obtained

446   using super-resolution imaging of mammalian cell chromatin structures where active

447   histone marks coincided with less compact chromatin and exhibited a higher degree of

448   co-localization with other active marks and RNAP II, while repressive marks coincide

449   with densely packed chromatin and are spatially distant from active RNAP II (84).

450   Wang et al. (85) also noted a decrease in chromatin fiber width when imaging was

451   conducted using nuclei prepared from a quiescent state compared with nuclei

452   undergoing active transcription. These data imply that in the quiescent state, chromatin

453   fibers were more tightly packed together. This was contrasted by chromatin fibers

454   imaged in the actively transcribing state, which appeared to be loosely entangled with

455   each other and form a relatively thick structure.

456 We do recognize several caveats in our analysis, which is based on calculations

457 assuming linear chromosomes while the KSHV genome is circular, a topology which

458 has implications for interpretation of contact frequencies. Terminal repeats (TR) that

459 occupy approximately one sixth of KSHV genome (e.g. 28 kb [0.8 kb x 35 copies]) were

460 also not considered in our analyses due to difficulties in mapping sequence reads from

461 this region. Further studies are therefore needed to confirm where and how a TR

462 interacts with a unique region, which would have significant impact in KSHV reactivation

463 through recruitment of the LANA protein complex (60). In addition, we employed a

464 single TAD prediction tool for our analysis; one that allows for 3D modeling. However, it

465 has been documented that TAD prediction algorithms may be very sensitive to data

466 resolution and normalization resulting in calls that vary greatly between tools in terms of

467 TAD number, size, and certain biological properties such as CTCF enrichment around

468 boundaries (86, 87). At the same time, we do have confidence in the 3D structural

469 model of the KSHV genome generated herein and the adequacy of the TADbit

470 approach, based on its assessment conducted by simulations on artificially generated

471 genomes, including a 1-Mb circular chromosome, and subsequent partial verification by

472 fluorescence *in situ* hybridization imaging (50).

473 In summary, this study established 3D KSHV genomic structure models of

474 episomes during latency and early lytic phase of reactivation and revealed the dynamics

475 of TAD structural changes by transcription activation. The 3D genomic structures

476 bestow another element to the effectiveness of herpesvirus gene regulation, a system

477 which may provide an opportunity to further study TAD regulation through the

478 intersection of viral and host proteins.

479

Data Availability

481 The data sets supporting the conclusions of this study are available in the National

482 Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO)

483 repository with the accession code GSE163695

484

Acknowledgements

489

Funding

497   **Figure 1. Normalized frequencies of genomic contact map during latency at 500**

498   **bp resolution. (A) KSHV genome contact maps in latent PEL cells**. Genomic

499   contacts detected at more than expected frequencies are depicted in red, while

500   underrepresented contacts between loci are presented in blue. Contact maps are

501   shown for latent BCBL-1, BC1 and BC-3 cells. **(B)** Genomic contacts detected in latent

502   *i*SLK/BAC16 cells are depicted as described in (A). A color range scale is listed and

503   KSHV ORF maps are shown beneath each contact map.

504

505   **Figure 2. Superimposition of select histone modifications and protein factor**

506   **binding sites on genomic contacts of the latent KSHV genome.** (A) The BCBL-1

507   contact map from Figure 1A is shown. Latent TRDs numbered 1-30 with their borders

508   (blue triangles) (from Fig. 3A/B) are listed below the contact map. (B). Alignment of

509   histone modifications H3K4Me3, H3K27Ac, H3K27me3, and binding sites of K-Rta,

510   CTCF, SMC1, POLR2A are shown. Control IgG for background control is also shown.

511   The sequence reads were mapped to the KSHV genome. IGV snapshots and a KSHV

512   genome map are shown. Numbers on the left-hand side of each track denotes the

513   height of the peak (e.g., read depth). **(C). Zoomed view of CTCF, SMC1, K-Rta, and**

514   **RNAPII (POLR2A) enrichment at select regions of the KSHV genome** (i) K4-PAN

515   RNA, (ii) ORFs 45-50 and (iii) repeat region. **(D). Positional correlations among K-**

516   **Rta, CTCF, and RNAPII recruitment sites.** Density plot showing average K-Rta ChIP-

517   seq signals within ±3 kb regions around the center of CTCF or PORLA2 peaks.

518   Heatmap shows K-Rta signals on CTCF or POLRA2 CUT&RUN peaks. *y*-axis is ranked

519   according to CTCF or PORLA2 enrichment in descending order.

520   **Figure 3. Dynamics of genomic contacts by reactivation in TREx-BCBL-1 cells.**

521   **(A) Interaction Frequency Matrix.** (a) TRDs were calculated and visualized with

522   TADbit, and genomic domains defined by frequencies of DNA contacts are marked with

523   black solid lines. (b) TAD border alignment during latency (upper) and during

524   reactivation (lower) are shown. Dark and light grey arches indicate TRDs with higher

525   and lower than expected intra-TRD interactions, respectively. TADbit border robustness

526   (from 1 to 10) is identified by a color gradient from blue to red. **(B) Contact Heatmap:**

527   **BCBL-1 reactivation vs. latency**. Contacts are color coded as described in Figure 1

528   (red, increased contacts; blue decreased contacts). TRDs called during reactivation

529   (Fig. 3A/B) numbered 1-14 with their borders (red triangles) are listed below the contact

530   map. KSHV ORF map is included in the bottom panel. **(C) TAD border analyses.**

531   (a)Insulation score (border strength) is plotted. Reactivation reduced insulation and

532   border strength to create larger transcription units. **(D)** Compartment analysis. 500 bp

533   bins were identified as in compartment A or B by calculating eigenvectors.

534

535   **Figure 4. KSHV 3D genomic structure modeling.** Frequencies of ligated fragments

536   were used to calculate distances between fragments at 2 kb resolution. The average of

537   8 different models were used to draw the figures. KSHV 3D genomic structure before

538   (left) and after (right) triggering reactivation are shown. Color gradation associated with

539   KSHV genomic position are depicted at the bottom of figures. The movies of the 3D

540   figures including individual 8 models in shadow are presented in Supplemental Figure

541   S2. The dotted line represents the terminal repeat (TR) region. Each sphere marks 2 kb

542   of linear genomic sequence.

543 **Supplemental Figure S1.**

544 Preparation of pAG-MNase for CUT&RUN. Protein A or A&G domain fused MNase

545 (pAG-MNAse) was expressed in *E.coli*, and purified with Ni-NTA beads (Invitrogen).

546 Purified proteins were subjected to SDS-PAGE and stained with Coomassie blue.

547 Recombinant protein sequence and purified proteins are shown. The pAG-MNase was

548 used for CUT&RUN studies described in this manuscript.

549

550 **Supplemental Figure S2.**

551 **(A)** Representative valid sequence pairs in CHi-C. Overall sequence reads with invalid

552 sequences were summarized in bar charts. One of three replicates of BCBL-1 samples

553 for both before (top) and after triggering reactivation (bottom) are shown. **(B)** Predicted

554 size distribution of KSHV genomic restriction fragments generated during CHi-C library

555 preparation.

556

557 **Supplemental Figure S3 (Movies).**

558 KSHV 3D genomic structure models for both latency and reactivated samples are

559 shown. Gray shadow indicates each model and color string 5' (blue) to 3' (red) are

560 generated based on average of 10 different models. Blue ball (PAN RNA region), Red

561 ball (K-Rta promoter region), and Yellow ball (K12 genomic region) were marked for

562 clarity.

563

564

565   1.   **Bechtel JT, Liang Y, Hvidding J, Ganem D.** 2003. Host range of Kaposi's sarcoma-
566   associated herpesvirus in cultured cells. J Virol **77:**6474-6481.

567   2.   **Vieira J, O'Hearn PM.** 2004. Use of the red fluorescent protein as a marker of Kaposi's
568   sarcoma-associated herpesvirus lytic gene expression. Virology **325:**225-240.

569   3.   **Johnson AS, Maronian N, Vieira J.** 2005. Activation of Kaposi's sarcoma-associated
570   herpesvirus lytic gene expression during epithelial differentiation. J Virol **79:**13769-
571   13777.

572   4.   **Arias C, Weisburd B, Stern-Ginossar N, Mercier A, Madrid AS, Bellare P, Holdorf M,**
573   **Weissman JS, Ganem D.** 2014. KSHV 2.0: a comprehensive annotation of the Kaposi's
574   sarcoma-associated herpesvirus genome using next-generation sequencing reveals
575   novel genomic and functional features. PLoS Pathog **10:**e1003847.

576   5.   **Bruce AG, Barcy S, DiMaio T, Gan E, Garrigues HJ, Lagunoff M, Rose TM.** 2017.
577   Quantitative Analysis of the KSHV Transcriptome Following Primary Infection of Blood
578   and Lymphatic Endothelial Cells. Pathogens **6**.

579   6.   **Renne R, Lagunoff M, Zhong W, Ganem D.** 1996. The size and conformation of Kaposi's
580   sarcoma-associated herpesvirus (human herpesvirus 8) DNA in infected cells and virions.
581   J Virol **70:**8151-8154.

582   7.   **Lagunoff M, Ganem D.** 1997. The structure and coding organization of the genomic
583   termini of Kaposi's sarcoma-associated herpesvirus. Virology **236:**147-154.

584   8.   **Toth Z, Brulois K, Lee HR, Izumiya Y, Tepper C, Kung HJ, Jung JU.** 2013. Biphasic
585   euchromatin-to-heterochromatin transition on the KSHV genome following de novo
586   infection. PLoS Pathog **9:**e1003813.

587   9.   **Toth Z, Brulois K, Jung JU.** 2013. The chromatin landscape of Kaposi's sarcoma-
588   associated herpesvirus. Viruses **5:**1346-1373.

589   10.   **Frohlich J, Grundhoff A.** 2020. Epigenetic control in Kaposi sarcoma-associated
590   herpesvirus infection and associated disease. Semin Immunopathol **42:**143-157.

591   11.   **Campbell M, Yang WS, Yeh WW, Kao CH, Chang PC.** 2020. Epigenetic Regulation of
592   Kaposi's Sarcoma-Associated Herpesvirus Latency. Front Microbiol **11:**850.

593   12.   **Chang PC, Fitzgerald LD, Hsia DA, Izumiya Y, Wu CY, Hsieh WP, Lin SF, Campbell M,**
594   **Lam KS, Luciw PA, Tepper CG, Kung HJ.** 2011. Histone demethylase JMJD2A regulates
595   Kaposi's sarcoma-associated herpesvirus replication and is targeted by a viral
596   transcriptional factor. J Virol **85:**3283-3293.

597   13.   **Campbell M, Watanabe T, Nakano K, Davis RR, Lyu Y, Tepper CG, Durbin-Johnson B,**
598   **Fujimuro M, Izumiya Y.** 2018. KSHV episomes reveal dynamic chromatin loop formation
599   with domain-specific gene regulation. Nat Commun **9:**49.

600   14.   **Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B.** 2012. Topological
601   domains in mammalian genomes identified by analysis of chromatin interactions.
602   Nature **485:**376-380.

603   15.   **Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum**
604   **NL, Meisig J, Sedat J, Gribnau J, Barillot E, Bluthgen N, Dekker J, Heard E.** 2012. Spatial
605   partitioning of the regulatory landscape of the X-inactivation centre. Nature **485:**381-
606   385.

607   16.   **Dekker J, Mirny L.** 2016. The 3D Genome as Moderator of Chromosomal
608   Communication. Cell **164:**1110-1121.

609   17.   **Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H,**
610   **Tanay A, Cavalli G.** 2012. Three-dimensional folding and functional organization
611   principles of the Drosophila genome. Cell **148:**458-472.
612   18.   **Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit**
613   **I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine**
614   **M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J.** 2009.
615   Comprehensive mapping of long-range interactions reveals folding principles of the
616   human genome. Science **326:**289-293.
617   19.   **Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A,**
618   **Lamb J, Nusbaum C, Green RD, Dekker J.** 2006. Chromosome Conformation Capture
619   Carbon Copy (5C): a massively parallel solution for mapping interactions between
620   genomic elements. Genome Res **16:**1299-1309.
621   20.   **Gomez-Marin C, Tena JJ, Acemel RD, Lopez-Mayorga M, Naranjo S, de la Calle-**
622   **Mustienes E, Maeso I, Beccari L, Aneas I, Vielmas E, Bovolenta P, Nobrega MA,**
623   **Carvajal J, Gomez-Skarmeta JL.** 2015. Evolutionary comparison reveals that diverging
624   CTCF sites are signatures of ancestral topological associating domains borders. Proc Natl
625   Acad Sci U S A **112:**7542-7547.
626   21.   **Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J,**
627   **Meyer BJ.** 2015. Condensin-driven remodelling of X chromosome topology during
628   dosage compensation. Nature **523:**240-244.
629   22.   **Hsieh TH, Weiner A, Lajoie B, Dekker J, Friedman N, Rando OJ.** 2015. Mapping
630   Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell **162:**108-119.
631   23.   **Mizuguchi T, Fudenberg G, Mehta S, Belton JM, Taneja N, Folco HD, FitzGerald P,**
632   **Dekker J, Mirny L, Barrowman J, Grewal SIS.** 2014. Cohesin-dependent globules and
633   heterochromatin shape 3D genome architecture in S. pombe. Nature **516:**432-435.
634   24.   **Le TB, Imakaev MV, Mirny LA, Laub MT.** 2013. High-resolution mapping of the spatial
635   organization of a bacterial chromosome. Science **342:**731-734.
636   25.   **Marbouty M, Le Gall A, Cattoni DI, Cournac A, Koh A, Fiche JB, Mozziconacci J, Murray**
637   **H, Koszul R, Nollmann M.** 2015. Condensin- and Replication-Mediated Bacterial
638   Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution
639   Imaging. Mol Cell **59:**588-602.
640   26.   **Wang X, Le TB, Lajoie BR, Dekker J, Laub MT, Rudner DZ.** 2015. Condensin promotes
641   the juxtaposition of DNA flanking its loading site in Bacillus subtilis. Genes Dev **29:**1661-
642   1675.
643   27.   **Rowley MJ, Corces VG.** 2016. The three-dimensional genome: principles and roles of
644   long-distance interactions. Curr Opin Cell Biol **40:**8-14.
645   28.   **Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM.** 2008. Cohesins
646   localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2
647   insulators. EMBO J **27:**654-666.
648   29.   **Kang H, Lieberman PM.** 2009. Cell cycle control of Kaposi's sarcoma-associated
649   herpesvirus latency transcription by CTCF-cohesin interactions. J Virol **83:**6199-6210.
650   30.   **Kang H, Lieberman PM.** 2011. Mechanism of glycyrrhizic acid inhibition of Kaposi's
651   sarcoma-associated herpesvirus: disruption of CTCF-cohesin-mediated RNA polymerase
652   II pausing and sister chromatid cohesion. J Virol **85:**11159-11169.

653 31. **Kang H, Wiedmer A, Yuan Y, Robertson E, Lieberman PM.** 2011. Coordination of KSHV
654   latent and lytic gene control by CTCF-cohesin mediated chromosome conformation.
655   PLoS Pathog **7:**e1002140.
656 32. **Chen HS, Wikramasinghe P, Showe L, Lieberman PM.** 2012. Cohesins repress Kaposi's
657   sarcoma-associated herpesvirus immediate early gene transcription during latency. J
658   Virol **86:**9454-9464.
659 33. **Kang H, Cho H, Sung GH, Lieberman PM.** 2013. CTCF regulates Kaposi's sarcoma-
660   associated herpesvirus latency transcription by nucleosome displacement and RNA
661   polymerase programming. J Virol **87:**1789-1799.
662 34. **Li DJ, Verma D, Mosbruger T, Swaminathan S.** 2014. CTCF and Rad21 act as host cell
663   restriction factors for Kaposi's sarcoma-associated herpesvirus (KSHV) lytic replication
664   by modulating viral gene transcription. PLoS Pathog **10:**e1003880.
665 35. **Toth Z, Smindak RJ, Papp B.** 2017. Inhibition of the lytic cycle of Kaposi's sarcoma-
666   associated herpesvirus by cohesin factors following de novo infection. Virology **512:**25-
667   33.
668 36. **Li D, Mosbruger T, Verma D, Swaminathan S.** 2020. Complex Interactions between
669   Cohesin and CTCF in Regulation of Kaposi's Sarcoma-Associated Herpesvirus Lytic
670   Transcription. J Virol **94**.
671 37. **Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT,**
672   **Hookway TA, Guo C, Sun Y, Bland MJ, Wagstaff W, Dalton S, McDevitt TC, Sen R,**
673   **Dekker J, Taylor J, Corces VG.** 2013. Architectural protein subclasses shape 3D
674   organization of genomes during lineage commitment. Cell **153:**1281-1295.
675 38. **Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL,**
676   **Machol I, Omer AD, Lander ES, Aiden EL.** 2014. A 3D map of the human genome at
677   kilobase resolution reveals principles of chromatin looping. Cell **159:**1665-1680.
678 39. **Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S.**
679   2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain
680   architecture. Cell Rep **10:**1297-1309.
681 40. **Toth Z, Maglinte DT, Lee SH, Lee HR, Wong LY, Brulois KF, Lee S, Buckley JD, Laird PW,**
682   **Marquez VE, Jung JU.** 2010. Epigenetic analysis of KSHV latent and lytic genomes. PLoS
683   Pathog **6:**e1001013.
684 41. **Gunther T, Grundhoff A.** 2010. The epigenetic landscape of latent Kaposi sarcoma-
685   associated herpesvirus genomes. PLoS Pathog **6:**e1000935.
686 42. **Kim KY, Huerta SB, Izumiya C, Wang DH, Martinez A, Shevchenko B, Kung HJ, Campbell**
687   **M, Izumiya Y.** 2013. Kaposi's sarcoma-associated herpesvirus (KSHV) latency-associated
688   nuclear antigen regulates the KSHV epigenome by association with the histone
689   demethylase KDM3A. J Virol **87:**6782-6793.
690 43. **Hu J, Yang Y, Turner PC, Jain V, McIntyre LM, Renne R.** 2014. LANA binds to multiple
691   active viral and cellular promoters and associates with the H3K4methyltransferase
692   hSET1 complex. PLoS Pathog **10:**e1004240.
693 44. **Belaghzal H, Dekker J, Gibcus JH.** 2017. Hi-C 2.0: An optimized Hi-C procedure for high-
694   resolution genome-wide mapping of chromosome conformation. Methods **123:**56-65.
695 45. **Schmitt AD, Hu M, Ren B.** 2016. Genome-wide mapping and analysis of chromosome
696   architecture. Nat Rev Mol Cell Biol **17:**743-755.

697 46. **Schoenfelder S, Javierre BM, Furlan-Magaril M, Wingett SW, Fraser P.** 2018. Promoter
698 Capture Hi-C: High-resolution, Genome-wide Profiling of Promoter Interactions. J Vis Exp
699 doi:10.3791/57320.
700 47. **Jung I, Schmitt A, Diao Y, Lee AJ, Liu T, Yang D, Tan C, Eom J, Chan M, Chee S, Chiang Z,**
701 **Kim C, Masliah E, Barr CL, Li B, Kuan S, Kim D, Ren B.** 2019. A compendium of
702 promoter-centered long-range chromatin interactions in the human genome. Nat Genet
703 **51:**1442-1449.
704 48. **Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot**
705 **E.** 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome
706 Biol **16:**259.
707 49. **Serra F, Bau D, Goodstadt M, Castillo D, Filion GJ, Marti-Renom MA.** 2017. Automatic
708 analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly
709 chromatin colors. PLoS Comput Biol **13:**e1005665.
710 50. **Trussart M, Serra F, Bau D, Junier I, Serrano L, Marti-Renom MA.** 2015. Assessing the
711 limits of restraint-based 3D modeling of genomes and genomic domains. Nucleic Acids
712 Res **43:**3465-3477.
713 51. **Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE.**
714 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J
715 Comput Chem **25:**1605-1612.
716 52. **Skene PJ, Henikoff S.** 2017. An efficient targeted nuclease strategy for high-resolution
717 mapping of DNA binding sites. Elife **6**.
718 53. **Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, Andrews S.**
719 2015. HiCUP: pipeline for mapping and processing Hi-C data. F1000Res **4:**1310.
720 54. **Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL.** 2016.
721 Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom.
722 Cell Syst **3:**99-101.
723 55. **Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H,**
724 **Glass CK.** 2010. Simple combinations of lineage-determining transcription factors prime
725 cis-regulatory elements required for macrophage and B cell identities. Mol Cell **38:**576-
726 589.
727 56. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nat
728 Methods **9:**357-359.
729 57. **Kim D, Paggi JM, Park C, Bennett C, Salzberg SL.** 2019. Graph-based genome alignment
730 and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol **37:**907-915.
731 58. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers**
732 **RM, Brown M, Li W, Liu XS.** 2008. Model-based analysis of ChIP-Seq (MACS). Genome
733 Biol **9:**R137.
734 59. **Papp B, Motlagh N, Smindak RJ, Jin Jang S, Sharma A, Alonso JD, Toth Z.** 2019.
735 Genome-Wide Identification of Direct RTA Targets Reveals Key Host Factors for Kaposi's
736 Sarcoma-Associated Herpesvirus Lytic Reactivation. J Virol **93**.
737 60. **Kumar A, Lyu Y, Yanagihashi Y, Chantarasrivong C, Majerciak V, Salemi M, Wang KH,**
738 **Inagaki T, Chuang F, R.R. D, Tepper CG, Nakano K, Izumiya C, Shimoda M, Nakajima KI,**
739 **Merleev A, Zheng ZM, Campbell M, Izumiya Y.** 2022. KSHV Episome Tethering Sites on

740   Host Chromosomes and Regulation of Latency-Lytic Switch by CHD4. Cell Reports **In**
741   **press**.

742 61. **Freese NH, Norris DC, Loraine AE.** 2016. Integrated genome browser: visual analytics
743   platform for genomics. Bioinformatics **32:**2089-2095.

744 62. **Shen L, Shao N, Liu X, Nestler E.** 2014. ngs.plot: Quick mining and visualization of next-
745   generation sequencing data by integrating genomic databases. BMC Genomics **15:**284.

746 63. **Davis RR, Campbell M, Izumiya Y, Tepper CG.** 2020. Capture Hi-C: Characterization of
747   chromatin contacts, vol 18. Elsevier Inc.

748 64. **Gavrilov A, Eivazova E, Priozhkova I, Lipinski M, Razin S, Vassetzky Y.** 2009.
749   Chromosome conformation capture (from 3C to 5C) and its ChIP-based modification.
750   Methods Mol Biol **567:**171-188.

751 65. **Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AI, Bochkov ID,**
752   **Chinnappan D, Cutkosky A, Li J, Geeting KP, Gnirke A, Melnikov A, McKenna D,**
753   **Stamenova EK, Lander ES, Aiden EL.** 2015. Chromatin extrusion explains key features of
754   loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U
755   S A **112:**E6456-6465.

756 66. **Bertero A, Fields PA, Ramani V, Bonora G, Yardimci GG, Reinecke H, Pabon L, Noble**
757   **WS, Shendure J, Murry CE.** 2019. Dynamics of genome reorganization during human
758   cardiogenesis reveal an RBM20-dependent splicing factory. Nat Commun **10:**1538.

759 67. **Schifano JM, Corcoran K, Kelkar H, Dittmer DP.** 2017. Expression of the Antisense-to-
760   Latency Transcript Long Noncoding RNA in Kaposi's Sarcoma-Associated Herpesvirus. J
761   Virol **91**.

762 68. **Qin J, Li W, Gao SJ, Lu C.** 2017. KSHV microRNAs: Tricks of the Devil. Trends Microbiol
763   **25:**648-661.

764 69. **Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W.** 2002. Looping and interaction
765   between hypersensitive sites in the active beta-globin locus. Mol Cell **10:**1453-1465.

766 70. **Li L, Zhang JA, Dose M, Kueh HY, Mosadeghi R, Gounari F, Rothenberg EV.** 2013. A far
767   downstream enhancer for murine Bcl11b controls its T-cell specific expression. Blood
768   **122:**902-911.

769 71. **Rubin AJ, Barajas BC, Furlan-Magaril M, Lopez-Pajares V, Mumbach MR, Howard I, Kim**
770   **DS, Boxer LD, Cairns J, Spivakov M, Wingett SW, Shi M, Zhao Z, Greenleaf WJ, Kundaje**
771   **A, Snyder M, Chang HY, Fraser P, Khavari PA.** 2017. Lineage-specific dynamic and pre-
772   established enhancer-promoter contacts cooperate in terminal differentiation. Nat
773   Genet **49:**1522-1528.

774 72. **Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen CA, Schmitt AD, Espinoza CA, Ren B.**
775   2013. A high-resolution map of the three-dimensional chromatin interactome in human
776   cells. Nature **503:**290-294.

777 73. **Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EE.**
778   2014. Enhancer loops appear stable during development and are associated with
779   paused polymerase. Nature **512:**96-100.

780 74. **Platt JL, Salama R, Smythies J, Choudhry H, Davies JO, Hughes JR, Ratcliffe PJ, Mole DR.**
781   2016. Capture-C reveals preformed chromatin interactions between HIF-binding sites
782   and distant promoters. EMBO Rep **17:**1410-1421.

75. **Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, Cavalli G.** 2017. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell **171:**557-572 e524.

76. **Stadhouders R, Vidal E, Serra F, Di Stefano B, Le Dily F, Quilez J, Gomez A, Collombet S, Berenguer C, Cuartero Y, Hecht J, Filion GJ, Beato M, Marti-Renom MA, Graf T.** 2018. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat Genet **50:**238-249.

77. **Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenkov VV, Ecker JR, Thomson JA, Ren B.** 2015. Chromatin architecture reorganization during stem cell differentiation. Nature **518:**331-336.

78. **Siersbaek R, Madsen JGS, Javierre BM, Nielsen R, Bagge EK, Cairns J, Wingett SW, Traynor S, Spivakov M, Fraser P, Mandrup S.** 2017. Dynamic Rewiring of Promoter-Anchored Chromatin Loops during Adipocyte Differentiation. Mol Cell **66:**420-435 e425.

79. **Criscione SW, De Cecco M, Siranosian B, Zhang Y, Kreiling JA, Sedivy JM, Neretti N.** 2016. Reorganization of chromosome architecture in replicative cellular senescence. Sci Adv **2:**e1500882.

80. **Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong CT, Cubenas-Potts C, Hu M, Lei EP, Bosco G, Qin ZS, Corces VG.** 2015. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Mol Cell **58:**216-231.

81. **Ray J, Munn PR, Vihervaara A, Lewis JJ, Ozer A, Danko CG, Lis JT.** 2019. Chromatin conformation remains stable upon extensive transcriptional changes driven by heat shock. Proc Natl Acad Sci U S A **116:**19431-19439.

82. **Shimoda M, Lyu Y, Wang KH, Kumar A, Miura H, Meckler JF, Davis RR, Chantarasrivong C, Izumiya C, Tepper CG, Nakajima KI, Tuscano J, Barisone G, Izumiya Y.** 2021. KSHV transactivator-derived small peptide traps coactivators to attenuate MYC and inhibits leukemia and lymphoma cell growth. Commun Biol **4:**1330.

83. **Gwack Y, Baek HJ, Nakamura H, Lee SH, Meisterernst M, Roeder RG, Jung JU.** 2003. Principal role of TRAP/mediator and SWI/SNF complexes in Kaposi's sarcoma-associated herpesvirus RTA-mediated lytic reactivation. Mol Cell Biol **23:**2055-2067.

84. **Xu J, Ma H, Jin J, Uttam S, Fu R, Huang Y, Liu Y.** 2018. Super-Resolution Imaging of Higher-Order Chromatin Structures at Different Epigenomic States in Single Mammalian Cells. Cell Rep **24:**873-882.

85. **Wang Y, Maharana S, Wang MD, Shivashankar GV.** 2014. Super-resolution microscopy reveals decondensed chromatin structure at transcription sites. Sci Rep **4:**4477.

86. **Dali R, Blanchette M.** 2017. A critical assessment of topologically associating domain prediction tools. Nucleic Acids Res **45:**2994-3005.

87. **Zufferey M, Tavernari D, Oricchio E, Ciriello G.** 2018. Comparison of computational methods for the identification of topologically associating domains. Genome Biol **19:**217.
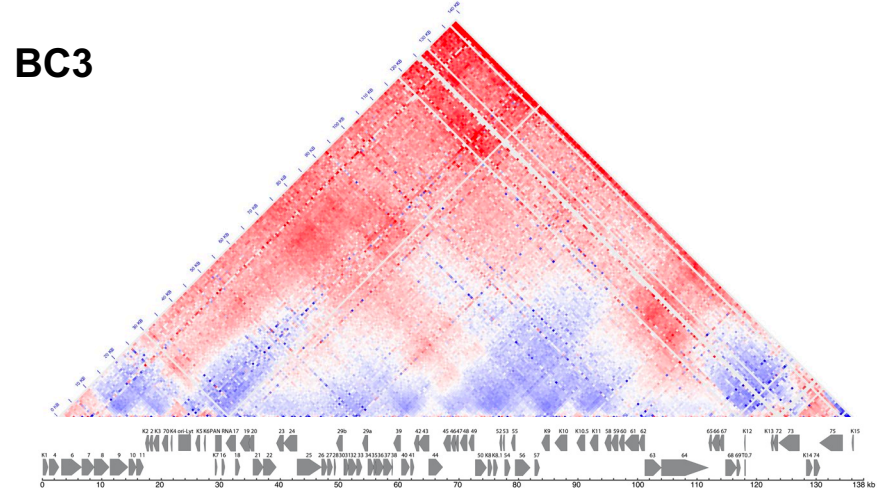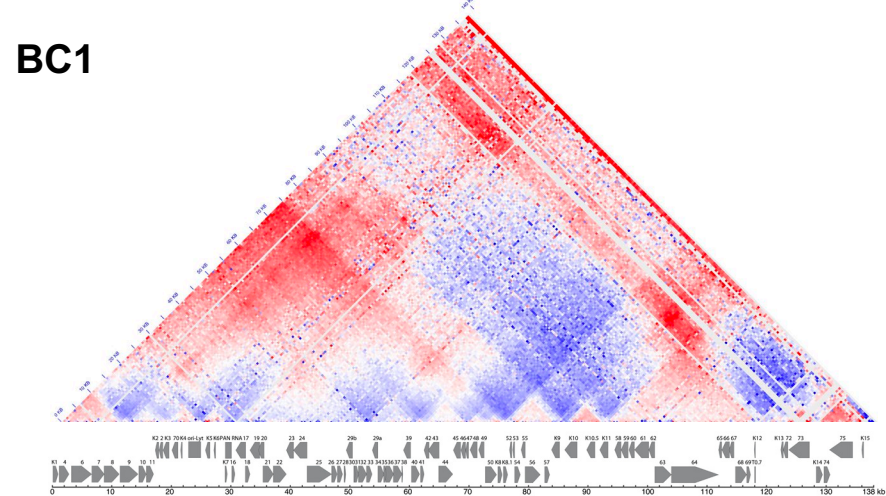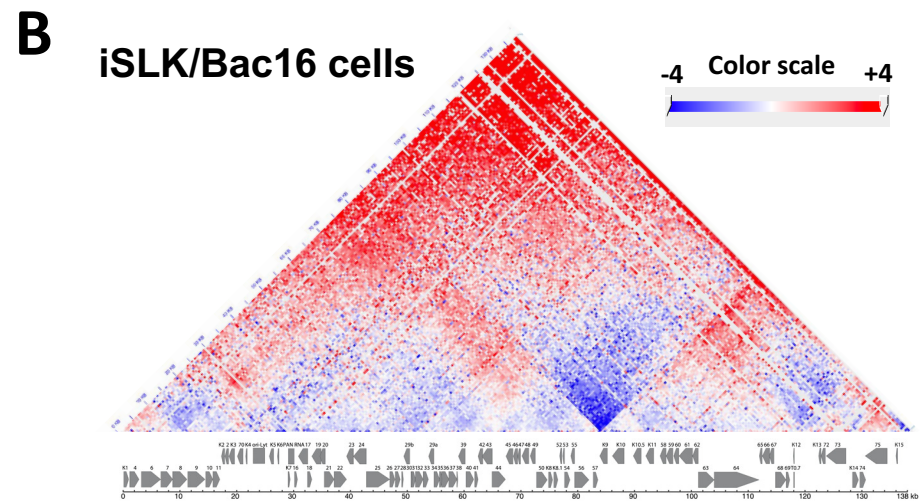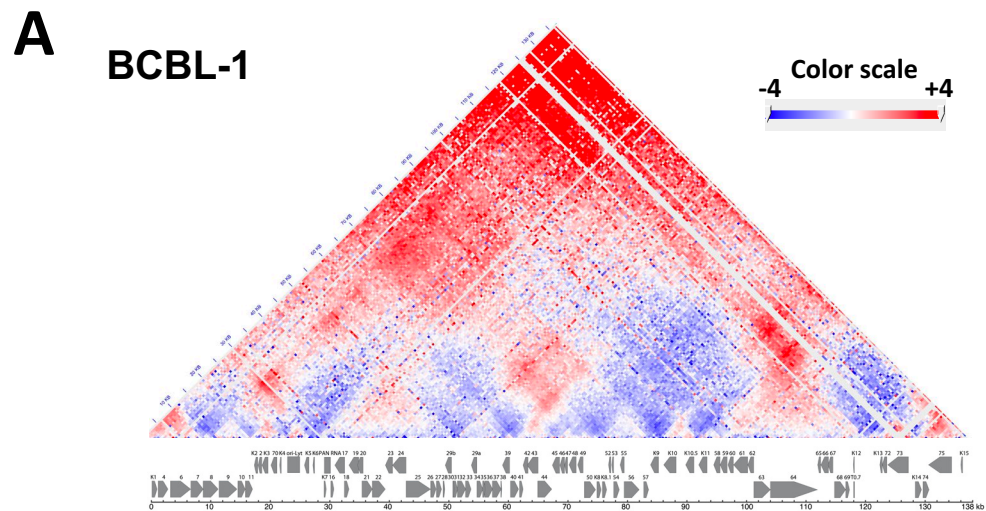
A

BCBL-1

Color scale −4 +4

BC1

BC3

B

iSLK/Bac16 cells

Color scale −4 +4

Figure 1

Figure 2

A

(a)

Latency

TRD number

During reactivation

Log2 Hi-C normalized interaction count

Resolution = 500 bp

(b)

Relative Hi-C count

Latency

Relative Hi-C count

Reactivated

Border scores
1 2 3 4 5 6 7 8 9 10

Resolution = 500 bp

B

-60 +60
color scale

Insulation score
Border strength

C

Latency

Reactivated (24 h)

Resolution = 500 bp

D

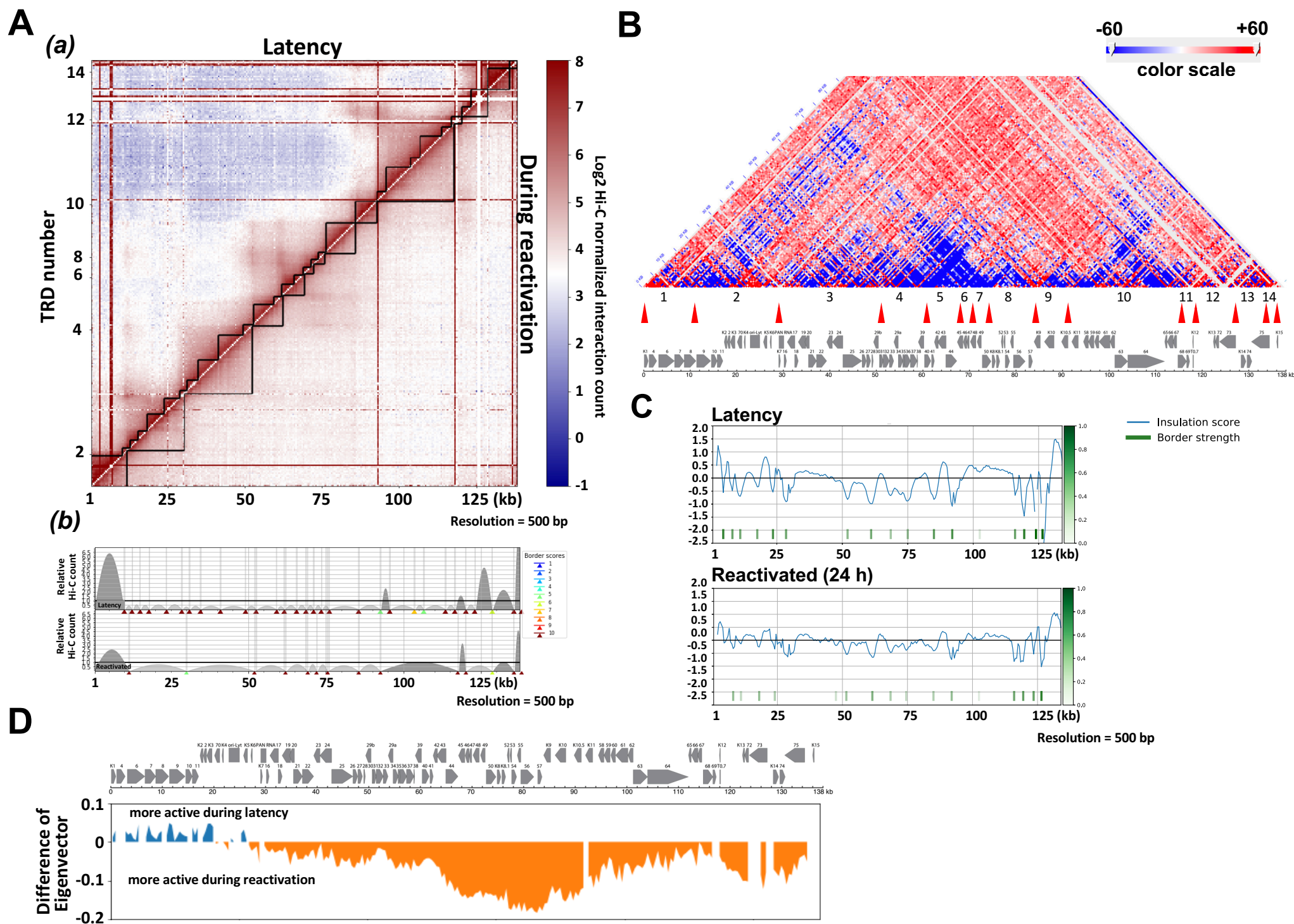Difference of Eigenvector

more active during latency

more active during reactivation

Figure 3

BCBL-1

Latency

Reactivation (24 hr)

90°

Figure 4