Vaccinia virus infection induces concurrent alterations in host chromatin architecture, accessibility, and gene expression

Vrinda Venu[1], Cullen Roth[2], Samantha H. Adikari[3], Eric M. Small[1], Shawn R. Starkenburg[2], Karissa Y. Sanbonmatsu[4,5], Christina R. Steadman[1]*

[1]Climate, Ecology & Environment Group, Los Alamos National Laboratory, Los Alamos, New Mexico

[2]Genomics & Bioanalytics Group, Los Alamos National Laboratory, Los Alamos, New Mexico

[3]Biochemistry & Biotechnology Group, Los Alamos National Laboratory, Los Alamos, New Mexico

[4]Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico

[5]New Mexico Consortium, Los Alamos New Mexico

*corresponding author: crsteadman@lanl.gov

Keywords: chromatin; structure; architecture; vaccinia virus; smallpox; nucleosome; accessibility; viral infection; three-dimensional genome; gene expression; epigenomics; Hi-C; ATAC-seq; RNA-seq

Abstract

Genomic DNA folds into complex configurations that produce particular local and global structures thought to profoundly impact genome function. To understand the dynamic nature of this relationship, we investigated the extent of host chromatin structural and functional changes in response to a viral agent. We performed comprehensive assessments of host architecture (Hi-C), accessibility (ATAC-seq), and gene expression (RNA-seq) in a paired manner in response to attenuated vaccinia (smallpox) virus. Over time, infection significantly increased long-range intra-chromosomal interactions and decreased chromatin accessibility. Fine-scale accessibility changes were independent of broad-scale chromatin compartment exchange, which increased (up to 12% of the genome) over time, underscoring potential independent mechanisms for global and local chromatin reorganization. The majority of differentially expressed genes, including those downregulated in immune responses, had concurrent alterations in local accessibility and loop domain restructuring. Increased B compartmentalization, intra-chromosomal interactions, and decreased inter-chromosomal interactions and chromatin accessibility together indicate that infection converts the host genome into a more condensed state with nearly equal bidirectional differential gene expression. These changes in host chromatin features may have implications for developing efficacious anti-viral countermeasures. Overall, our empirical data provides evidence of orchestrated concurrent alterations in chromatin architecture, accessibility, and gene expression in response to infection, further reinforcing the notion of coordinated structure-function dynamics of the genome.

Introduction

While genomic loci can be linearly very distant from one another, they fold into complex configurations in three-dimensional space to enable proximal interaction[1]. These configurations have hierarchical architecture, and along with chemical chromatin modifications, they influence—or are often influenced by—genome functions such as replication, recombination, and gene expression[2]. The flexible, self-organizing, fluctuating nature of the genome is thought to have a major role in its functional plasticity, and several studies have used polymer physics models of chromosomes to predict structure and functionality[3-7]. Recent work demonstrates that a complex, combinatorial, dynamic structure-function relationship exists between chromatin architecture and accessibility, the epigenetic landscape, and genome functionality[8-10]. Deeper empirical evidence from responsive cellular states is required to understand the causality and influence of global versus local chromatin architecture on function. Intrinsic and extrinsic cellular cues alter chromatin architecture, including shifts in proximity during the cell cycle, differentiation[11,12], and in response to infection, drug treatments, and environmental stressors[13-15]. Yet, detailed observations of concurrent alterations in global and local chromatin structure and gene function are still limited. To this end, we investigated the extent of changes in host chromatin interactions, accessibility, gene expression and their correlations using a viral infection model.

Viruses hijack host cellular function, including mechanisms involved in chromatin modification, to enhance viral replication, survival, and proliferation[16]. Studies have demonstrated that many double stranded DNA viruses (and some RNA viruses) that translocate into the nucleus directly access and interact with transcriptional machinery for replication and host chromosomes for interference of defense programs[17,18]. Non-nuclear viruses, including

RNA viruses such as SARS-CoV-2, may impart significant changes to the host chromatin, including altering gDNA methylation and histone modifications, and in some cases, chromatin structure[8,13,19]. These studies suggest that viral-induced changes to chromatin structure and alteration of host (and virus) epigenetic modifications aids in the progression of infection. However, detailed characterization of host chromatin structure-function dynamics due to non-nuclear viral infection has only been investigated in a handful of cases[8,13]. Given the diversity of infection mechanisms and immunogenicity among various pathogenic viruses, attenuated viral strains, and virus-like particles, a clear understanding of host chromatin dynamics and functional responses would greatly enhance our ability to create counter measures to infections and effective vaccine development. To meet this need, we established a virus infection model using an attenuated vaccinia virus strain to characterize host chromatin structure-function dynamics over the course of viral infection.

The virus family of Poxviridae is comprised of particularly resilient and pathogenic viruses that cause severe infectious disease, including smallpox. Effective eradication of smallpox occurred over forty years ago with an aggressive global vaccination program using vaccinia virus (VACV), a poxvirus with poorly understood origins but efficacious outcomes in humans[20]. The Poxviridae family is unique among dsDNA viruses: they complete their lifecycles entirely in the cytoplasm of host cells[21]. While their large genomic DNA replicates within the host cytoplasm, viral particle release requires protein localization in the host nucleus suggesting interaction with the host genome is required for virulence[22,23]. To establish an experimentally amenable viral infection model to permit interrogation of genomic structure-function relationships, we chose to infect a permissible mammalian cell line (Vero) with an attenuated modified vaccinia virus Ankara (MVA) strain. Vero cell lines, originally derived from African

4

green monkey (*Chlorocebus sabaeus*) kidney cells, are used in many vaccine production platforms[24]. To date, subsets of these cell lines have been considered permeable in varying degrees to VACV and its attenuated strains, including MVA[25,26]. While MVA is non-replicative in most mammalian cells, it expresses all classes of its genes upon cellular entry, synthesizes viral DNA, and produces immature virions that are not released[27]. MVA vectors are promising candidates for their ability to express transgenes and immunogenicity *in vivo*[25]. Several recombinant vaccines that employ MVA vectors have been developed against diseases including HIV[28], Zika virus[29], and SARS-CoV-2[30,31].

To obtain detailed, high resolution chromatin architecture data, we employed Hi-C, the most recent iteration of chromosome conformation capture sequencing technology, which can illuminate all contacts within a genome at up to 1 kb resolution[32]. Hi-C uses *in situ*, proximity-based ligation followed by sequencing for genome-wide characterization of 3D chromatin configurations. These include chromatin domains with preferential interactions and distinctive boundaries between actively transcribed euchromatin versus inactive transcriptionally repressed heterochromatin, or A and B compartments, respectively[33]. While simplistic, partitioning of the genome into A and B compartments occurs on a large (MB) scale, and genomic architectural rearrangements occur in at least 36% of the genome during differentiation, resulting in distinctive juxtaposition pattern of A and B compartments[11]. Recent studies have demonstrated that chromatin accessibility assays complement Hi-C contact maps, and *cis* proximal loci may have similar accessibility, suggesting that higher-order structure may influence function[34]. Mapping of chromatin accessibility may be performed using Assay for Transposase Accessible Chromatin by Sequencing (ATAC-seq) which provides information on nucleosome-free regions. The ease of this method underscores its relative ubiquity in the epigenomics field.

Here, we attempt to overcome the seemingly mutually exclusive nature of genome architecture variation and fine-tuned resolution analysis at the gene level by pairing Hi-C, ATAC-seq, and transcriptomics. Using these analyses, we investigated the extent of host chromatin structure and function changes in response to a viral agent without direct contact in the nucleus. We found that MVA infection alters global chromatin organization, including increased chromatin contact frequency and reduced chromatin accessibility as infection progressed over a 24-hour period. Though not linear, we detected close association of fine-scale chromatin accessibility with gene expression and chromatin loop formation. Changes in gene expression that involved downregulation of certain proinflammatory genes were concurrent with restructuring of chromatin loop domains and local accessibility reduction. These findings underscore the dynamic plasticity and correlation of both genome structure and function providing evidence that structure modulates but is not deterministic of function in response to viral infection.

Results

To investigate concurrent alterations in host chromatin accessibility and 3D organization as viral infection progresses, we infected a susceptible cell line (Vero), derived from African green monkey, with MVA (Figure 1A). Using immunofluorescence microscopy, we confirmed successful infection and its gradual progression at 12-, 18-, and 24-hours post infection (hpi) (Figure 1B). At each time point, cells were collected from mock-infected (control) and MVA-infected cell culture flasks for paired measurement of 3D genomic organization, chromatin accessibility, and gene expression using Hi-C, ATAC, and RNA sequencing, respectively (Figure 1A). We generated and sequenced 12 Hi-C libraries, with each replicate per condition totaling over 600 million 150-bp read pairs. We detected approximately 700 million valid Hi-C contacts per condition per time point (12, 18, and 24 hpi) (Table S1). Twelve ATAC libraries

6

and 12 RNA-seq libraries (both paired to the aforementioned Hi-C libraries) were generated and sequenced to around 40 million (Table S2) and 17 million read pairs per library, respectively (Table S3). We observed high replicate concordance in all three datasets (Figure S1); thus, replicates were combined for visualization. For all other analyses, replicates were kept separate unless specified otherwise. Principal component analysis of ATAC-seq read pileup, Hi-C contact frequency, and gene expression (Figures 1C, D, E, respectively) demonstrated distinct separation of mock-infected control cultures from MVA-infected cultures, indicating global changes in host chromatin architecture, accessibility, and expression due to viral infection. Using these high-resolution paired Hi-C, ATAC- and RNA-seq datasets with matching mock-infected controls at each time point, we report concurrent changes in chromatin architecture and accessibility as viral infection progressed in this model system and functional outcomes of MVA infection.

MVA infection reorganizes host chromatin architecture

Analysis of Hi-C sequencing data revealed substantial alteration of chromatin contacts in response to MVA infection. For example, chromosome 27 displayed some of the most significant changes in chromatin architecture as visualized via Hi-C contact maps (Figure 2A) and contact ratio maps (Figure 2B). Contact maps compare mock-infected controls (lower diagonal) to MVA-infected cells (upper diagonal) at each time point (Figure 2A). Fold change profiles of contacts (Figure 2B) display increased intensity (red color) in the off-diagonal region and the lack of color on the diagonal proximal region indicating more alterations in long-range intra-chromosomal contacts compared to short-range contacts.

The healthy Vero genome is clearly organized into areas of chromatin activity denoted as A (active) or B (inactive) compartments. As viral infection progressed, the number of genomic
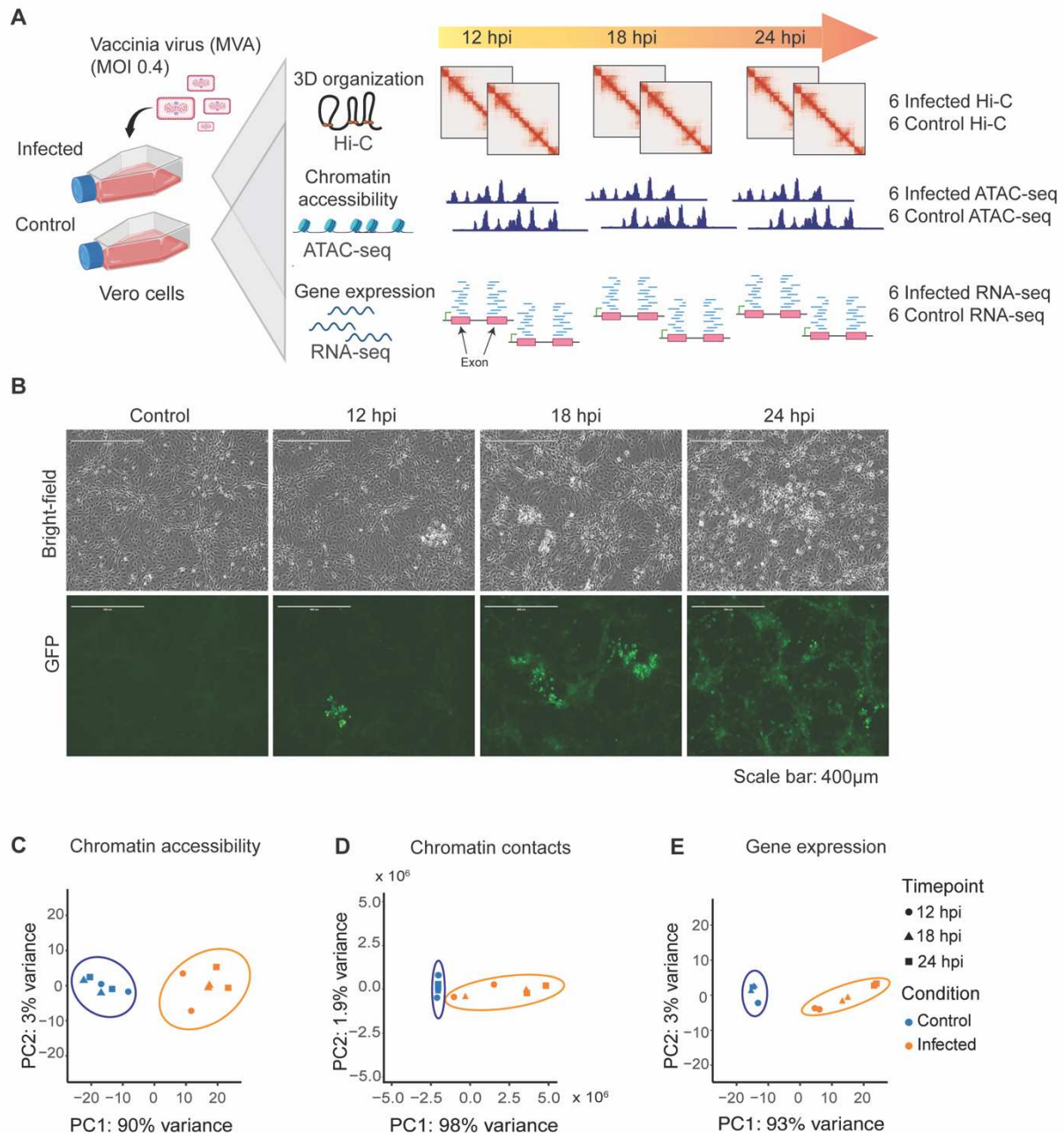
Figure 1: Infection model system reveals changes in chromatin architecture, accessibility, and gene expression over time. *(A) Schematic representation of the experimental design where Modified Vaccinia virus Ankara (MVA) was used to infect Vero cells with an MOI of 0.40. Vero cells cultivated in the same flask were divided for Hi-C, ATAC-seq, and RNA-seq experiments to provide paired data sets for direct*

comparison. Cells from mock-infected (control) and MVA-infected cultures were collected at 12-,18- and 24-hours post infection (hpi) with two biological replicates each. A total of 12 Hi-C libraries, 12 ATAC-seq libraries, and 12 RNA-seq libraries were prepared. (B) Infection efficiency was verified by immunofluorescence using an anti-vaccinia virus monoclonal antibody (green). Cell density and morphology can be seen in the bright-field images. (C) Principal component analysis of genome-wide chromatin accessibility from ATAC-seq data, (D) chromatin contacts from Hi-C sequencing data, E) gene expression levels from RNA-seq data demonstrate clear separation between mock-infected control and MVA-infected cultures.

regions exchanging compartments increased indicating global changes in genome organization. Compartment assignments were calculated using a Pearson correlation coefficient for mock-infected control and MVA-infected Hi-C maps (Figure 2C). Representative large-scale compartment exchanges are visualized at 100 kb resolution, where highlighted areas (red boxes) indicate regions of exchange as infection progressed (Figure 2C). The strength of this compartment exchange, indicative of complete versus weak exchange, increased as infection progressed. To quantify compartment exchange, we identified co-variate regions in Hi-C data at 100 kb resolution using principal component analysis. Using ATAC-seq data, we assigned co-variate regions as activate (A) or inactive (B) compartments based on high or low accessibility, respectively. We found the number of regions strongly exchanging compartments (A to B, B to A) considerably increased as infection progressed (Figure 2D). For all contacts assessed, 2.4% of the Vero genome exchanged compartments at 12 hpi, 7.8% at 18 hpi, and 12% at 24 hpi, with a slight bias toward A to B exchange (Figure 2D). Compartment exchange occurs in response to a number of biological processes, including viral infection; here, these findings suggest that MVA infection moderately alters the compartment profile of the Vero genome into a less 'active' mode with greater propensity toward B compartmentalization.
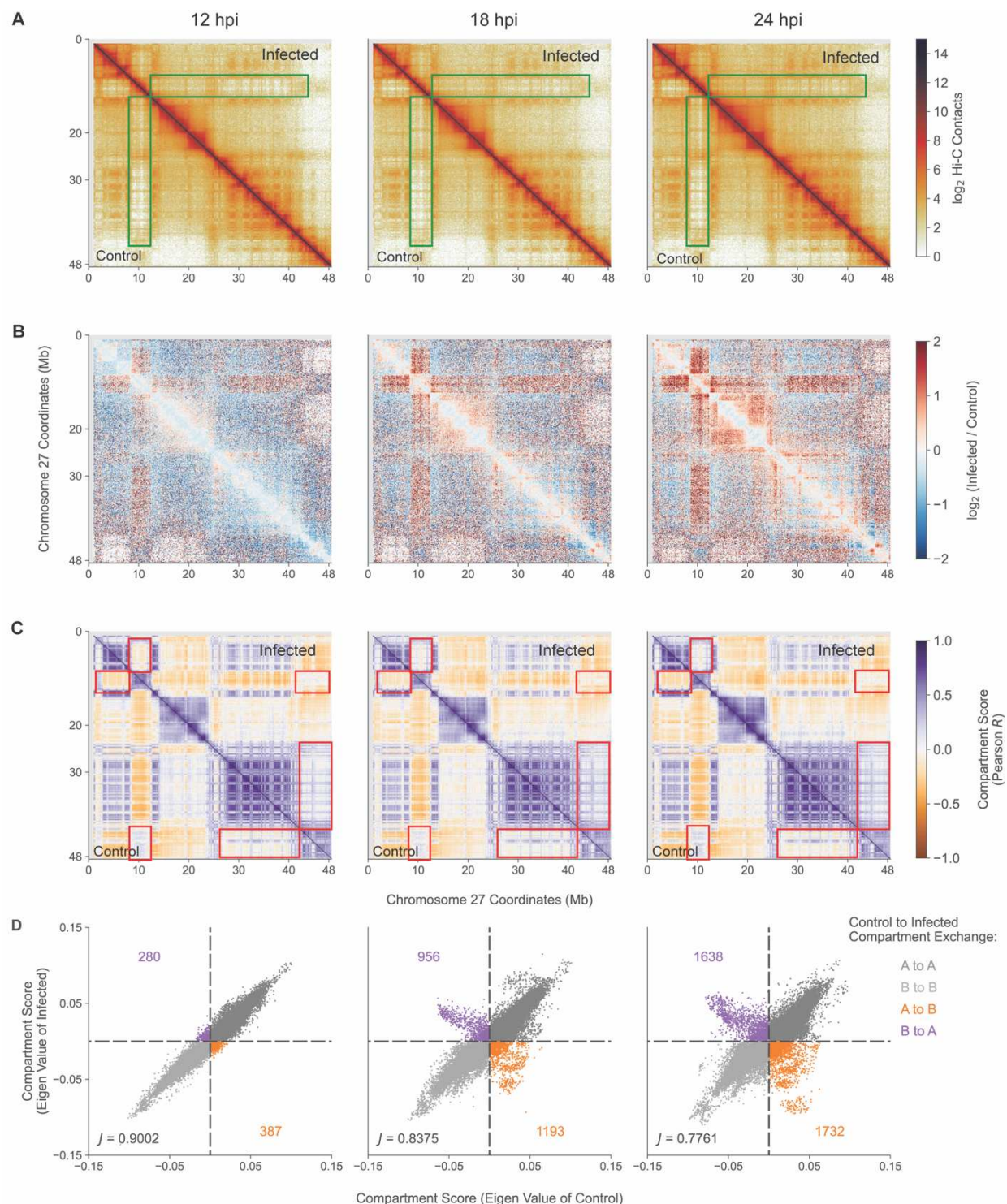
Figure 2: Modified Vaccinia virus Ankara (MVA) infection induced progressive changes in Vero host chromatin 3D architecture. A) Hi-C contact matrices of chromosome 27 in mock-infected control (lower diagonal region) versus MVA-infected (upper diagonal region) cells at 12-, 18-, and 24 hours post

infection (hpi). Green boxes indicate regions with pronounced contact differences between control and MVA-infected cells, where bin size is 100 kb. (B) Differential (infected/control) contact matrices of chromosome 27 across all three time points are shown at 100 kb resolution. Blue indicates control-biased contacts, where more contacts were observed in control cells, and red indicates infection-biased contacts, where more contacts were observed in MVA-infected cells. (C) Pearson correlation matrix of chromosome 27 in mock-infected control (lower diagonal region) versus MVA-infected (upper diagonal region) cells, where bin size is 100 kb. Red boxes highlight the regions with altered compartmentalization. (D) Scatterplots demonstrating replicate averaged compartment scores (eigen value) of control vs infected cells in 100 kb bins across the genome at all three time points. Off-diagonal points represents the compartment exchange events, where orange indicates A to B compartment exchange and purple indicates B to A compartment exchange. Counts of compartment exchange are annotated along the off diagonal. The Jaccard index is annotated in the bottom left and measures the overlap between compartment assignments between mock-infected control and MVA-infected scores.

Infection increased mid- to long-range intra-chromosomal interactions

To further understand the changes in 3D genome organization, we assessed the frequency of contacts among chromatin from Hi-C data, which provides a direct measure of spatial interaction between any two points in the genome at a given time[35]. From this analysis, we found that mid- to long-range chromosome interactions increased as MVA infection progressed across 12, 18 and 24 hpi time points (Figure 3A). This was determined from plotting the contact frequency as a function of the distance between regions in contact, where we observed the greatest difference in the number of the contacts in the 500 kb to 5 Mb range with a bias toward increased contacts in MVA-infected cells. This difference is clearly visualized from the contact frequency fold change of the genomic distance (Figure 3A; bottom panel). We corroborated this finding using the multiHiCcompare[36] tool, which normalizes contact frequency in 100 kb bins across each

chromosome and detects infection- and control-biased contact regions (Figure 3B). We denote 'infection-biased' bins as regions where MVA infection increased chromatin interactions more frequently than interactions observed in mock-infected control cells; similarly, 'control-biased' bins represent regions where MVA infection induced a loss of chromatin interactions compared to control cells. From this broad scale analysis, there were more interactions in the mid- to long-range distance (500 kb to 5 Mb, light green colored points in Figure 3B) in MVA-infected cells compared to controls. We further observed a significant increase in long distance contacts as infection progressed. At 24 hpi, approximately 66% (479) of the identified significantly different bins (726) represented mid- to long-range distance interactions. However, in a fine-scale analysis (in 10 kb bins across each chromosome), less than 28% (68) of the significantly different bins represent mid- to long-range distance interactions, whereas 71% of the bins (169) represented short-range distance (< 500 kb) interactions (Figure S2A). Relatively fewer bins with significant contact differences were observed across the time course in both mock-infected control and MVA-infected cells (Figure S2 B, C. Only broad-scale analysis result is shown). These analyses demonstrate that intra-chromosomal contacts were increased after viral infection across the entire Vero genome. Furthermore, the pattern of inter-chromosomal contacts was dependent upon individual chromosome behavior, where increased interactions were observed with specific combinations of chromosomes. These included interactions among chromosomes 4, 10, 19, 21, and 26 with 31; 9 and 19; 22 and 26 (Figure S3). However, there was a decrease in contacts among most chromosome pairs (Figure S3) at 24 hpi. This observation is contrary to that reported for SARS-CoV-2 infection of A549 cells[13], suggesting that MVA infection increased compactness of each chromosome potentially resulting in shrinkage of chromosomal territory and thereby weakening inter-chromosomal interactions. We speculate that the observed

predominantly mid- to long-range intra-chromosomal interaction changes likely represent sub-compartment repositioning, whereas the relatively low abundant short range interaction changes represent loop reorganization.

Chromatin loops co-localize regulatory elements with their genomic targets to bring one-dimensionally distal chromatin regions in proximity to one another via loop extrusion mechanisms[37]. These loop structures can be identified from Hi-C contact maps as concentrated contact points marking loop anchors, where the two segments of chromatin are co-localized. We identified approximately 14,000 loops per sample by combining loops called at five, ten, and 25 kb resolution (Table S4). Viral infection did not have a significant impact on loop structure compared to controls at early time points. This was determined from differential loop analysis comparing the contact frequency within loop anchor points between MVA-infected and mock-infected control cells at each time point. 'Infection-biased' loops were defined as those with more contacts in MVA-infected cells than in mock-infected controls: indeed, zero, 16, and 2097 infection-biased loops were identified at 12, 18 and 24 hpi, respectively. However, we did observe some loss of loops due to MVA infection (1136 control-biased loops at 24 hpi) (Figure 3C). This may indicate moderate fine-scale chromatin remodeling with prolonged infection. However, of the chromatin loops identified across the time course, infection-biased differential loops represent less than three percent, while control-biased loops constitute less than two percent of all the loops called. No significant differences in loops were detected among mock-infected control cells across time points or among MVA-infected cells across time points (Figure S4).
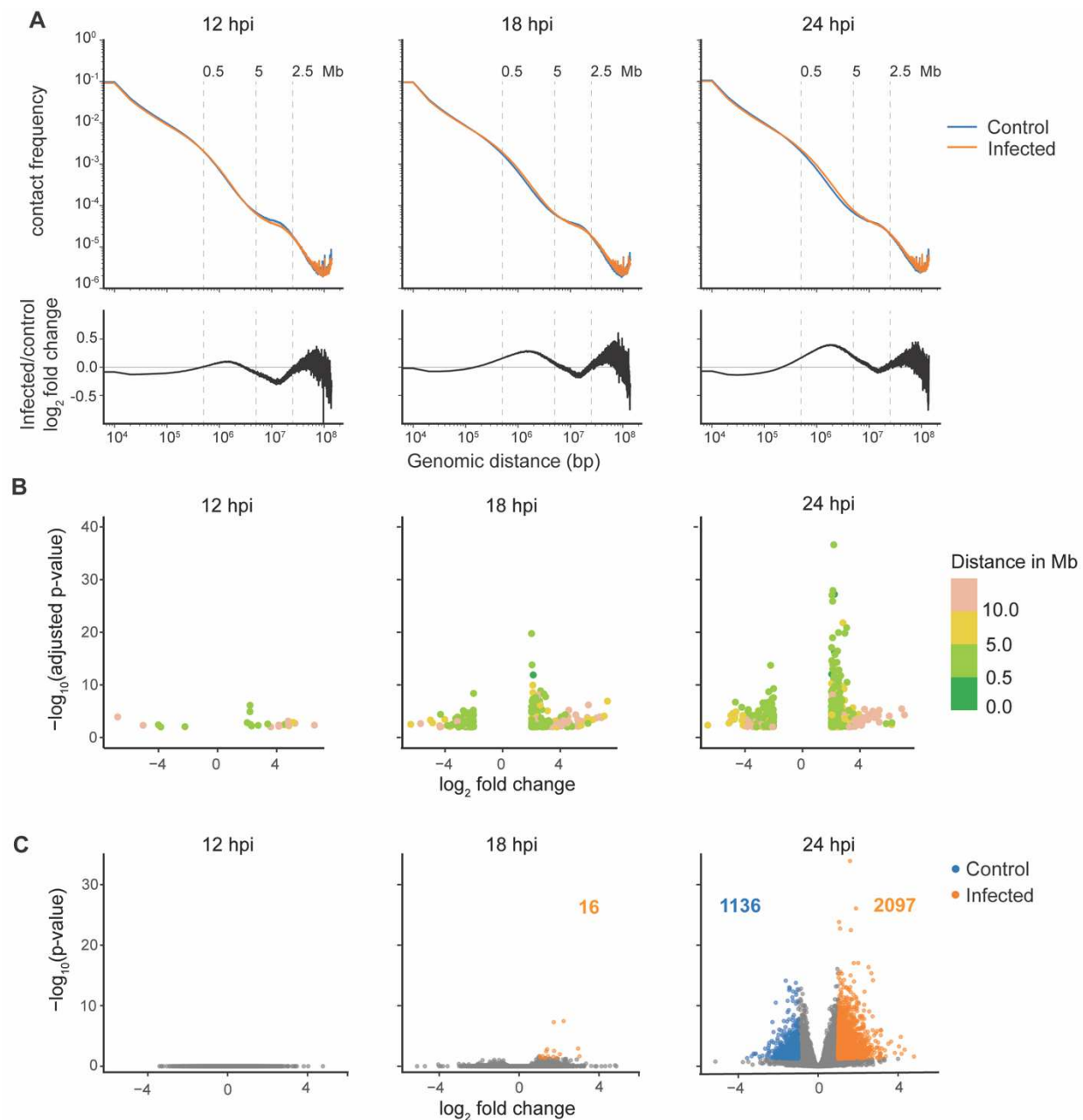
Figure 3: MVA infection predominantly alters mid- to long-range chromatin interactions. *(A) Contact frequency in control (blue) and MVA-infected (orange) cells as a function of genomic distance in log scale is plotted for all three time points; replicates are combined. Bottom panel: log2 fold change in contact frequency between control and infected cells is plotted as a function of genomic distance in log scale. (B) Genomic regions with significant differences in the number of contacts between mock-infected control and MVA-infected cultures were identified using multiHiCcompare at 100 kb resolution and*

plotted as log$_2$ fold change in contact frequency (x-axis) and negative log$_{10}$ adjusted p-value (y-axis). Color scale annotates the distance between contacting regions. Points in the positive axis represent infection-biased regions (more contacts due to MVA-infection than in control) and points in the negative axis represent control-biased regions. (C) Volcano plots represent identified differential loops between control and MVA-infected cells at all three time points. Grey dots represent all called differential loops where significant (adjusted p-value <0.05) control-biased loops (log$_2$ fold change <-1) are shown in blue and infection-biased loops (log$_2$ fold change >1) are shown in orange.

Infection reduced chromatin accessibility genome wide

After determining the extent of host chromatin 3D reorganization in response to MVA infection, we next investigated the fine-scale changes in chromatin accessibility across the genome using the paired ATAC-seq datasets generated from the same cell cultures used for the Hi-C experiments. A total of 300,667 distinct open chromatin regions (OCRs) spanning approximately 5.1% of the host genome were identified. MVA infection notably reduced global accessibility compared to mock-infected controls at 12, 18, and 24 hpi as demonstrated in heat maps (Figure 4A). Further, as infection progressed from 12 to 24 hrs, global accessibility considerably decreased in MVA-infected cells. Differential OCRs increased as infection progressed as well, determined by comparing read coverage at each individual OCR between mock-infected control and MVA-infected cells (Figure 4B). At each time point, differential OCRs were dominated by regions with reduced accessibility in MVA-infected cells, particularly from 12 to 18 hpi (Figure 4B) concurrent with reduced global accessibility. Further, significant accessibility changes were also detected from 12 to 24 hpi in MVA-infected cells and were dominated by OCRs with a reduction in their accessibility (Figure S5). This analysis suggests that the OCRs found in both conditions did not significantly change across the time points; rather, the accumulation or loss of

OCRs contributed to the differential accessibility we observed. Collectively, global host chromatin accessibility was reduced as MVA infection progressed in Vero cells.

To determine if the fine-scale accessibility changes were correlated with large-scale compartment exchange, we determined the overlap between differential OCRs with our previously identified A and B compartments. The majority of the differential OCRs (~80.3%) in MVA-infected cells overlapped with the compartments that remained active (A), and a substantial number (~8.4%) overlapped with compartments that remained inactive (B) (Figure 4C). While the greatest number of differential OCRs were identified at 24 hpi, approximately 10.5% of control-biased OCRs and 7% of infection-biased OCRs overlapped with exchanging compartments (Figure 4C). Regions with increased accessibility in control cultures are denoted as 'control-biased' OCRs and regions with increased accessibility in MVA-infected cultures as 'infection-biased' OCRs. Of the small fraction of differential OCRs that overlapped exchanging compartments, the majority of control-biased OCRs in MVA-infected cells were within compartments that exchanged from active (A) to inactive (B) compartments. This suggests that these open regions became less accessible and simultaneously moved to B compartments. Conversely, infection-biased OCRs in MVA-infected cells were within compartments that exchanged from inactive (B) to active (A) compartments (Figure 4C). Overall, the global accessibility change due to MVA infection was primarily independent of large-scale compartment exchange, except for less than 10% of the differential OCRs.

Next, we investigated whether chromatin accessibility change was associated with fine-scale chromatin organizational features such as chromatin loops by examining the ATAC-seq read coverage across differential loop anchor points (Figure 4D). We found correlations between loop formation and accessibility in both MVA-infected and mock-infected control cells: MVA-

16

infected cells had reduced accessibility at the anchor points of weakened loops (i.e., control-biased loops) (Figure 4D, left panel). Conversely, accessibility was retained to the same extent as controls at anchor regions of loops that were gained after MVA infection (Figure 4D, right panel). Thus, chromatin accessibility is potentially a pre-requisite for loop anchoring; however, loop re-organization does not seem to alter accessibility of an otherwise open region. This assertion is based on evidence that control cells retained accessibility at infection-biased loop anchor points, which had more contacts in MVA-infected cells than in controls (as depicted in schematic in the bottom panel of Figure 4D).

Given that a relationship exists between accessibility and expression[38-40], we sought to predict gene regulatory pathways that were potentially impacted by MVA infection. We focused on strong OCRs (annotated specifically to gene promoters) and categorized the genes as either 'activated' or 'suppressed' by MVA infection based on the overlap of the promoter with an infection-biased differential peak or a control-biased differential peak, respectively. We used KOBAS[41] gene enrichment analysis and found enrichment of several signaling pathways in the 'activated' category in mock-infected control cells, suggesting potentially 'suppressed' outcomes in MVA- infected cells. These pathways included pro-inflammatory NF-κB signaling pathway, TLR signaling pathway, and cytokine-cytokine interaction pathway, among others. Conversely, we identified pathways in the 'activated' category in MVA-infected cells suggesting potential upregulation that included various cell-to-cell signaling pathways such as Rap1 signaling pathway, Ras signaling pathway, and MAPK signaling pathway (Figure S6, S7).
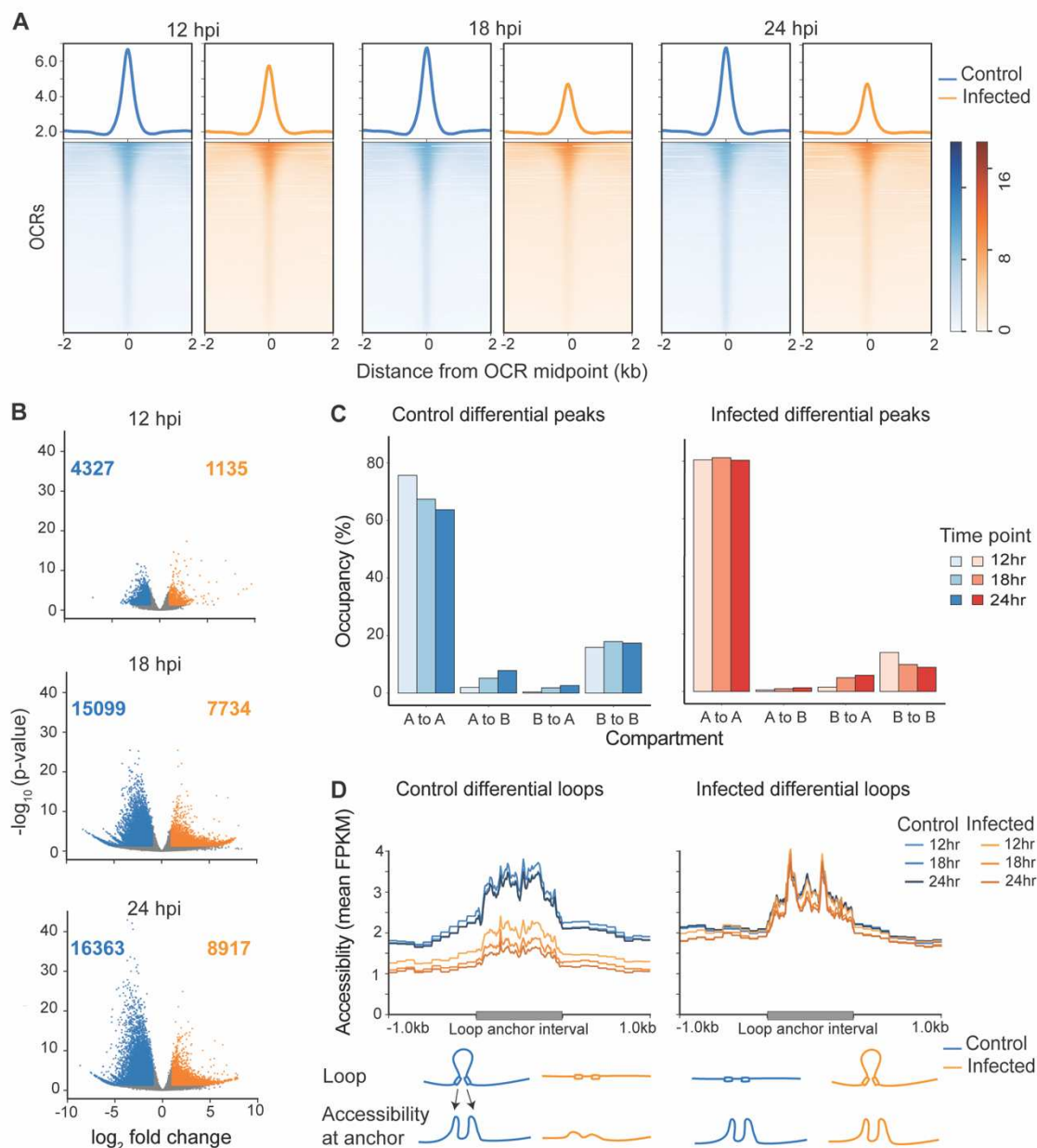
Figure 4: Global decrease in chromatin accessibility. *(A) Chromatin accessibility across consensus open chromatin regions (OCRs) is plotted for each time point 12-, 18-, and 24-hours post infection (hpi), where blue indicates mock-infected control cells and orange indicates MVA-infected cells. The ATAC-seq reads from all biological replicates are combined. The normalized average read count profile is plotted in the top panel, while the bottom panel shows the normalized read count per OCR where each row represents*

one consensus OCR. Read counts were determined at ± 2 kb region around all consensus OCRs. (B) Volcano plots demonstrate differentially accessible OCRs between mock-infected control and MVA-infected cells at each time point, 12, 18 and 24 hpi. Grey dots represent all OCRs. Blue dots indicate significantly different (adjusted p-value<0.05) control-biased OCRs (log$_2$ fold change< -1), while orange dots indicate infection-biased OCRs (log$_2$ fold change >1). Control-biased OCRs are regions that are accessible in control cells, while infection-biased OCRs are regions that are accessible in infected cells. (C) Bar plots depict the percentage overlap between mock-infected control (left) and MVA-infected (right) biased differential OCRs and various compartment categories at all three time points. (D) Time- and condition-wise chromatin accessibility (replicate combined normalized ATAC-seq read pileup) across control-biased (left) and infection-biased (right) loop anchor intervals ± 1 kb is plotted in the top panel. A schematic describing the observed relationship between loop anchor points and accessibility is shown in the bottom panel.

MVA infection altered gene expression correlated with chromatin accessibility

To obtain deeper insight into the functional relationship between chromatin architecture and accessibility, we quantified functional changes that occurred in response to MVA infection using paired RNA-seq datasets generated from the same cell cultures used for Hi-C and ATAC-seq experiments. As expected, we observed more genes with altered expression in MVA-infected cultures compared to mock-infected control cells as infection progressed (Figure 5A). A greater number of genes were upregulated in MVA-infected cells than in mock-infected controls at all three timepoints (12, 18, 24 hpi), with 836 genes with increased expression at 24 hpi. We used standard threshold of fold change and *p*-value (log$_2$ fold change > 2, p_{adj}<0.05) for this analysis, but when we considered all genes with significant differential expression, irrespective of their fold change, we found 4146 downregulated genes and 3959 upregulated genes out of 32,054 total genes at 24 hpi. Thus, there was a marginal increase in the number of downregulated genes at all

19

three time points, suggesting that MVA infection induced changes in host gene expression in both directions with the greatest differential expression, approximately 25% of genes, at 24 hpi.

Though chromatin accessibility change is not always concordant with changes in transcription[42], they are often associated with one another, either as a permissive factor or an effect of permissiveness[40]. Therefore, in the ATAC-seq and RNA-seq paired datasets, we investigated the chromatin accessibility around the transcription start sites (TSS) of significantly ($p_{adj}<0.05$) differentially regulated genes at 24 hpi, where the fold change was greater than two. MVA-infected cells had considerably lower accessibility at the promoters of genes that were downregulated compared to mock-infected control cells (Figure 5B). Conversely, changes in chromatin accessibility around the TSS of upregulated genes were marginally different in MVA-infected cells compared to control cells. This suggests that either increased accessibility is not a prerequisite for increased gene expression or increased gene expression does not elicit a significant change in accessibility. However, decreased gene expression is strongly correlated with decreased accessibility in response to MVA infection.

To determine if alterations in chromatin accessibility are a proxy for functional changes, we investigated whether genes with differential OCRs in their promoters had significantly different gene expression (Figure 5C). Only OCRs from merged bins of differential ATAC-seq peaks with a base mean greater than 100 and a fold change greater than one localized ± 500 bp around TSS were used. At 24 hpi, the majority of the genes with infection-biased accessibility (increased accessibility compared to controls) had increased gene expression in MVA-infected cells (orange filled circles in Figure 5C). Control-biased accessibility (reduced accessibility compared to mock-infected controls) had decreased gene expression in MVA-infected cells (filled blue circles in Figure 5C). Both of these findings suggest a distinct relationship between

20

accessibility and expression; however, we also detected upregulated genes with decreased accessibility (blue open circles in Figure 5C) and downregulated genes with increased accessibility (orange open circles in Figure 5C). Thus, the relationship between accessibility and gene expression, while strong, is not demonstratively linear. This finding corroborates other recent studies reporting that changes in accessibility cannot directly predict changes in gene expression; however, different categories of genes, including housekeeping genes, may have different relationships with accessibility and underscore the variability in this relationship.
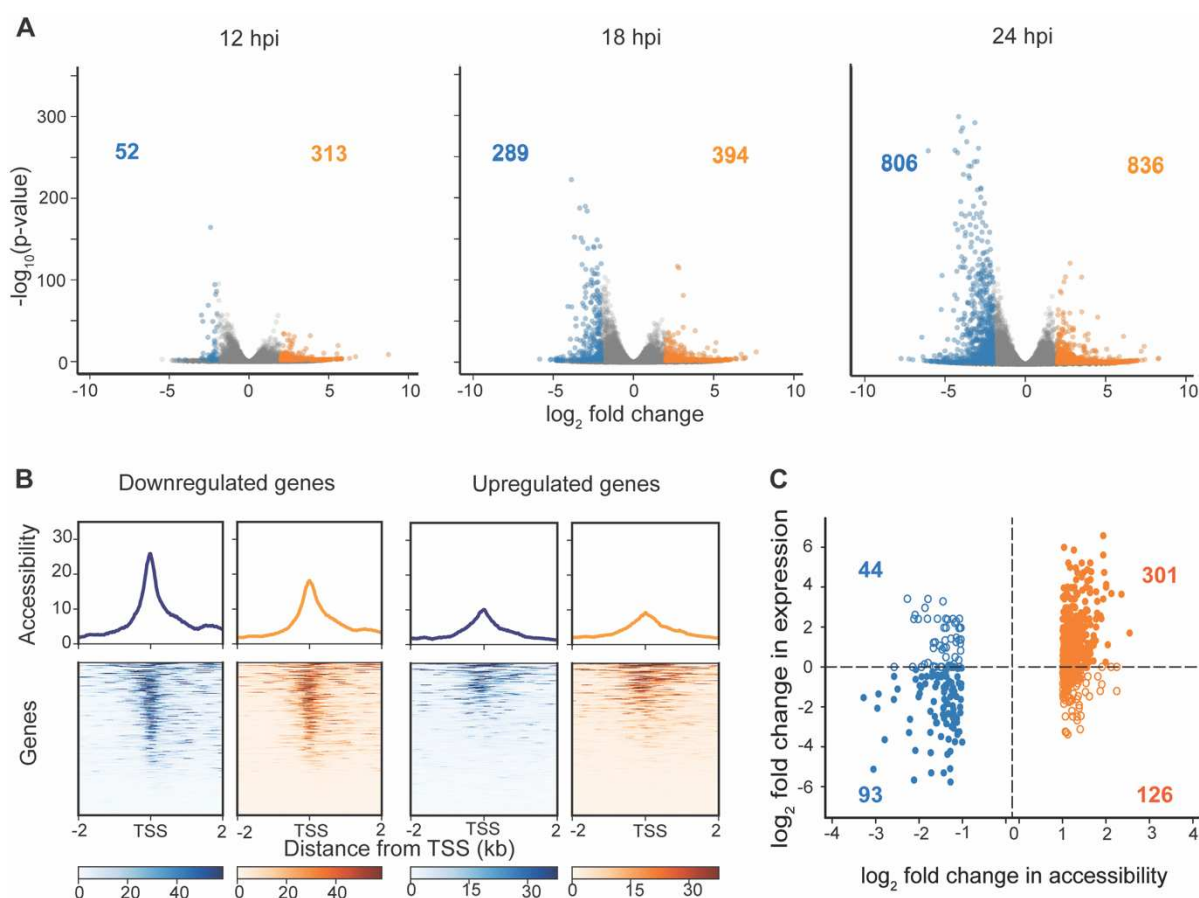

Figure 5: MVA infection induced gene expression changes and moderately correlates with local chromatin accessibility. A) Volcano plots demonstrate differentially expressed gene between mock-infected control and MVA-infected Vero cells at each time point: 12, 18 and 24 hpi. Grey dots represent

all genes. Blue dots indicate significantly different (adjusted p-value<0.05) down regulated genes (log$_2$ fold change <-2), while orange dots indicate significantly upregulated genes (log$_2$ fold change >2). B) Chromatin accessibility around TSS (± 2000bp) of strongly downregulated (left two panels) and upregulated (right two panels) genes at 24 hpi is plotted. Blue indicates accessibility data from mock-infected control Vero cells and orange indicates MVA-infected cells. The ATAC-seq reads from all biological replicates at 24 hpi are combined. The normalized average read count profile is plotted in the top panel, while the bottom panel shows the normalized read count around TSS where each row represents one consensus gene. C) Log$_2$ fold change in accessibility at promoter annotated OCRs (OCRs overlapping with TSS± 500 bp) vs log$_2$ fold change in expression in corresponding genes is plotted. OCRs accessible in control cells are shown in the negative x-axis (blue colored) and OCRs accessible in MVA-infected cells are shown in the positive x-axis (orange colored). Positive fold change in gene expression indicates upregulated genes in infected cells and vice versa.

Functional response correlates with architecture and accessibility

Gene expression ultimately dictates how a host cell responds to viral infection. After we identified differential gene expression (both up and downregulation), we investigated the regulatory pathways that may be activated or suppressed in response to infection. We focused on significant differential genes with an adjusted *p*-value less than 0.05 and a log$_2$ fold change greater than two and performed an enrichment analysis using the KOBAS[41] platform. This analysis identified activated pathways from upregulated genes and suppressed pathways from downregulated genes in MVA-infected cells. (The top 10 pathways are shown in Figure 6A; see Figure S8 and S9 for the extended list). MVA infection resulted in downregulation of genes within several pathways associated with cell proliferation including the TGFβ pathway, Hippo pathway (which includes WNT-related genes), and the PI3K-Akt pathway, all of which are heavily influenced by viral infection[43-45]. Pathways involving upregulated genes included those

22

associated with carbohydrate metabolism and insulin secretion, both of which are associated with the Warburg effect, an expected outcome of viral infection[46].

The chromatin accessibility profiles of several genes with differential expression displayed altered accessibility as infection progressed. For example, the promoter region of ZNF467 had increased accessibility (infection-biased OCRs) and associated increased gene expression (Figure 6B). Similarly, the promoter region of OTOP2 and USH1G (nearby genes on sense and antisense strands) with increased accessibility (infection-biased OCRs) had concurrently modest increased gene expression (Figure 6C). Genes that overlapped with control-biased OCRs (regions with reduced accessibility due to MVA infection) included IL1R1 (Figure 5D), LFNG, and TTYH3 genes (Figure 5E), all of which were downregulated in MVA-infected cells. These genes are associated with cellular immune responses, including the progression of viral infection involving the O-glycan pathway, which was also predicted in the ATAC-seq KOBAS enrichment analysis. Closer inspection of chromatin contacts around these genes suggested chromatin de-regulation due to infection; conversely, loss of transcription may have led to intra-domain condensation. For example, in the regions containing the IL1R1 (Figure 5D), LFNG, and TTYH3 genes (Figure 5E), we found an increased mixing of loop domains (i.e., the entire extruded region of a loop) with increased contact frequency and an associated decrease in accessibility in MVA-infected cells. Using a more generalized analysis, we found that gained/restructured loop domains in MVA-infected cells were larger in size than lost/weakened loops (median size of 400 kb instead of 200 kb in control cells). Furthermore, 64.7% of strong OCRs with increased accessibility and 72% of strong OCRs with reduced accessibility in 24 hpi cells overlapped with re-structured loop domains (empirical p-value <0.001 with 1.21- and 1.41-times increased association than expected by chance respectively). Along the same lines, we find

that 67.2% of upregulated genes and 70.3% of downregulated genes overlapped with re-structured loop domains (empirical p-value <0.001 with 1.38- and 1.22-times increased association than expected by chance respectively). These findings suggest a connection between gene expression changes and loop restructuring induced by MVA infection. Collectively, as infection progressed, differential gene expression was associated with concurrent alterations in 3D chromatin architecture and accessibility.

Discussion

The architecture of the genome, including its compartmentalization, chromatin contacts, conformation, and accessibility, is as fundamental to understanding the function of a particular genomic loci as its sequence. Indeed, investigation of 3D chromatin structure may be paramount to deciphering disease-states, aberrant cellular behavior, and the impact of environmental perturbations, including pathogen exposures, all of which have functional implications, including dysregulation of gene expression[13,47-50]. Further, pathogen-induced changes in host chromatin features may have far-reaching implications for development of appropriate and efficacious countermeasures including vaccine development. Although it is apparent that genome architecture is highly dynamic, the correlation between its structural organization at various scales and functional regulation is still unclear and requires comprehensive empirical data from various perturbation experiments. To this end, we employed a model system of infection for time course chromatin assessments in a paired manner allowing for direct comparison of global and local architectural features and functional changes.
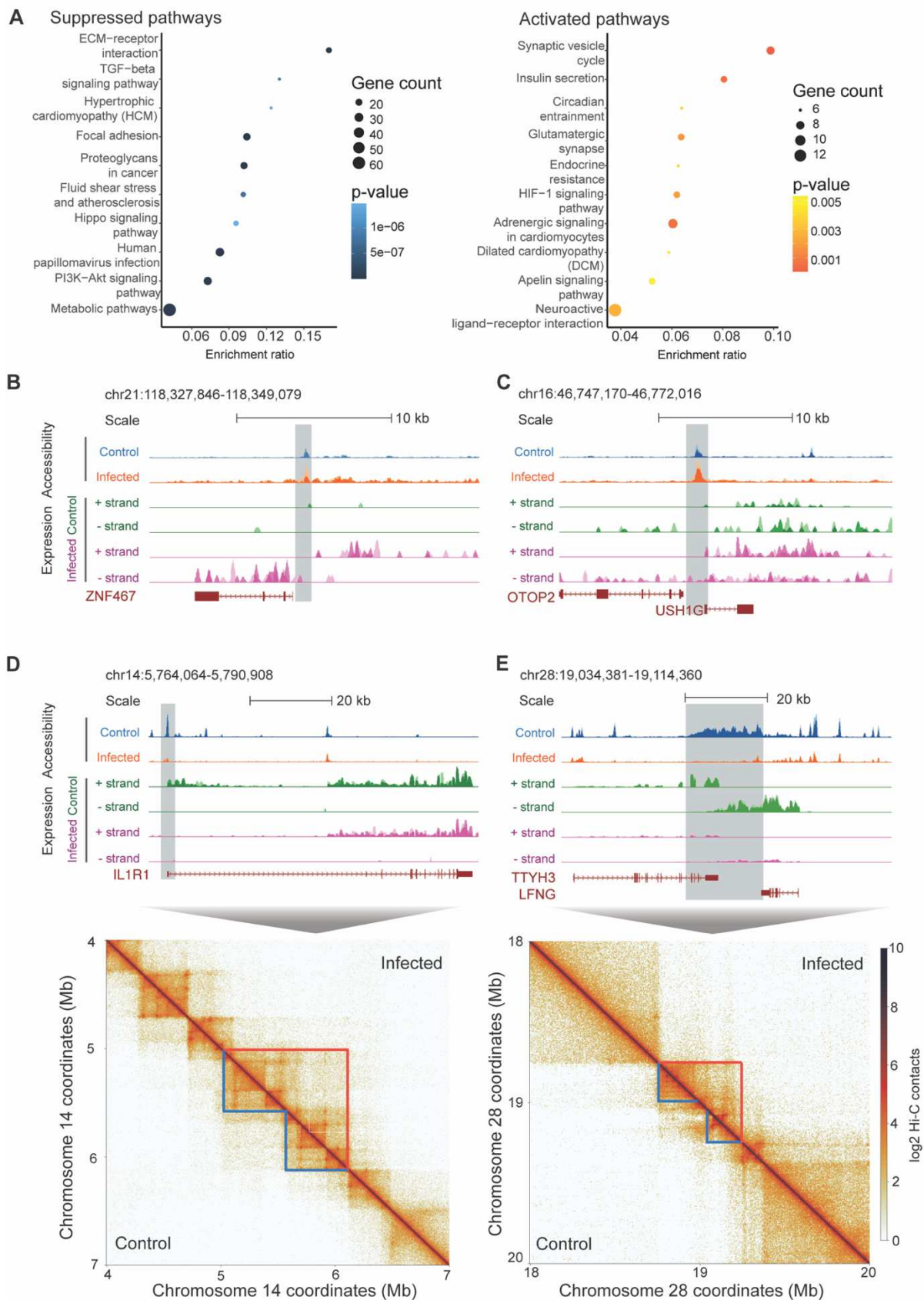
Figure 6: Functional responses correlate with changes in chromatin architecture and accessibility. (A) Top gene regulatory pathways that are affected by MVA infection predicted by Kegg Orthology analysis of differentially expressed genes. Top 10 suppressed pathways (left panel) and activated pathways (right panel) are shown in the decreasing order of their gene enrichment ratio. (B), (C) ATAC-seq and gene expression signal tracks from 24 hpi at example regions with infection-biased accessibility and gene expression are shown. Replicate signals are overlaid. Blue: accessibility in mock-infected control cells, Orange: accessibility in MVA-infected cells, Green: control gene expression, Purple: infected gene expression. (D), (E) ATAC-seq and gene expression signal tracks from 24 hpi at example regions with control-biased accessibility and gene expression are shown in the top panel (color scheme same as B and C). Corresponding Hi-C contact heat map of a larger region including the corresponding differential accessible window is shown in the bottom panel. Control and MVA-infected contacts are shown in the lower and upper diagonal parts of the figure, respectively. Blue and orange markings highlight contact domains with increased mixing in MVA-infected cells.

We found substantial changes in global and local chromatin architecture in MVA-infected Vero host genomes. There was an overall increase in mid- to long-range intra-chromosomal interactions in the host genome and prolonged infection increased changes in contact frequency. We speculate these viral-induced changes in mid- to long-range intra-chromosomal interactions likely represent sub-compartment repositioning. We detected a small fraction of intra-chromosomal interaction changes that localized to fine-scale loop anchors, marking differential loops after infection. Thus, prolonged infection caused moderate fine-scale chromatin remodeling. In fact, gained/strengthened loops due to infection were larger in size compared to lost/weakened loops suggesting that chromatin loop domain restructuring may propagate to larger chromatin architectural features. However, our findings suggest that gene expression may facilitate restructuring through accessibility alterations, though the directionality

of this interaction was not determined. While MVA conducts a non-permissive lifecycle in host cells[22], vaccinia-derived proteins evade innate immunity and suppress antiviral pathways[51]. Therefore, the observed alterations in Vero host genomic structure and function are part of the host defense response, which is either evoked or molecularly hijacked by the virus.

Concurrent with the increased intra-chromosomal interactions, 12% of the Vero genome exchanged compartments with a slight bias from activate (A) to inactive (B) conversion. In many organisms, the extent of compartment exchange is based on cellular phenotype and biological processes; for example, during cellular differentiation in human cells, approximately 36% of the genome strongly exchanges compartments[11], while 60% of the genome is dynamically compartmentalized between various cell types and tissues[52]. In response to the cytoplasmic RNA virus SARS-CoV-2, 30% of the human host genome (carcinomic lung epithelial cells, A549) either exchanges or weakens compartments[13]. In comparison, our findings suggest that MVA infection moderately converts the host Vero genome into a less 'active' mode over time, with greater propensity toward B compartmentalization. Increased contact frequency coupled with relatively few short-range interaction changes suggests that infection induced chromatin condensation due to newly formed larger loops with more contacts, which may reduce accessibility in the nuclear space.

Indeed, at the fine scale, infection reduced accessibility over time across the Vero genome. At any given point, we determined only ~5.1% of the Vero genome is accessible regardless of treatment; less than nine percent of these accessible regions is differentially accessible due to infection (at 24 hrs). While this represents a small fraction of the genome, our integrated analysis of chromatin accessibility with differential chromatin loops and differential gene expression suggests that accessibility change at anchor intervals and promoters is correlated

27

with the ability to form loops and modulate gene expression levels, respectively. Gene downregulation was strongly correlated with reduced promoter accessibility more so than upregulation was with increased accessibility. Further, reduced accessibility was observed at anchor points of loops that were lost during infection; yet infected cells retained the same amount of accessibility (as controls) in loops that were gained. While we did not perform causality studies to conclude whether loss of accessibility precludes loop loss, our results demonstrate that lack of a loop does not cause loss of accessibility, but loss of accessibility is correlated with loop loss. Therefore, we speculate that anchor point accessibility is a predetermining factor for loop formation. The functional relevance of chromatin loop re-organization is predicated on whether loops provide proximal interactions for genes with their enhancer or repressor elements. Further, reorganization may also control localization of genomic elements in the nuclear space thereby causing fine-scale functional changes or vice versa. In agreement with this assertion, we found that differential accessible regions and differentially expressed genes overlapped with restructured loop domains more than expected by chance.

At the functional level, direct examination of gene expression showed downregulation of many genes associated with pathways including host immunity and proliferation in response to viral infection, whereas those associated with cell signaling and carbohydrate metabolism were upregulated; these hallmark expression changes in response to viral infection have been demonstrated by many other studies[53,54]. In addition, our paired data revealed concurrent alteration in chromatin accessibility and contacts at several of those loci, particularly at genes that were strongly downregulated. Several reports provide evidence of viral hijacking of host epigenetic modifications, which have profound influence on gene expression and chromatin contacts; as such, MVA-induced alterations in epigenetic modifications may underscore the

changes we observe in gene expression which influence accessibility and chromatin architecture. Characterization of epigenetic changes was not within the scope of this study; however, future studies including comprehensive epigenome profiling would greatly advance our understanding. Further, drug or CRISPR-mediated global or locus-specific perturbation of transcription and/or epigenetic modifications would provide more insights into the directionality of this relationship.

While correlated dynamics were observed among chromatin accessibility, gene expression, and chromatin loops, differentially accessible regions and expressed genes only marginally overlapped with exchanging compartments. We would have anticipated that if the same factors responsible for large-scale compartment exchange were driving fine-scale accessibility changes as well, then accessibility would follow changes in compartmentalization. On the contrary, our finding indicates that fine-scale changes in nucleosome positioning in response to infection is likely independent of broad-scale chromatin compartmentalization mechanisms. Taken together, increased B compartmentalization, increased intra-chromosomal interactions with decreased inter-chromosomal interactions, and reduced chromatin accessibility suggest that MVA infection converts the host genome into a more condensed state. However, overall structural condensation did not correspond to an overall downregulation of gene expression resulting from MVA infection; rather, changes in gene expression occurred nearly equally in both directions. Potentially, stronger directional responses may be elicited by more virulent pathogens. These empirical observations support the hypothesis that genome structure does not determine function in a discrete manner, but rather, promotes and supports function[55]. Thus, Louis Sullivan's famously quipped '*form follows function*' not only applies to late 19th century architecture, but also to the impact of genome function on structure further underscoring the complexity of this relationship.

Methods

Cell culture and seeding for infection

Cryopreserved Vero cells, derived from the kidney tissue of a normal adult African green monkey, were sourced from the ATCC (Cat # CCL-81). Cells were thawed, passaged, expanded, and seeded for viral infection in complete medium consisting of Dulbecco's Modified Eagle's Medium (DMEM) containing 4 mM L-Glutamine (ATCC, Cat # 30-2002), supplemented with 10% (v/v) Fetal Bovine Serum (FBS, ATCC, Cat # 30-2020) and filter-sterilized before use. Vero cells were maintained at 37°C with 5% CO_2 and 95% relative humidity. Vero cells were cryopreserved at passage #7 from the original source vial using freeze medium consisting of DMEM supplemented with 10% (v/v) FBS and 5% (v/v) dimethylsulfoxide (DMSO, ATCC, Cat # 4-X).

Cryopreserved Vero cells were thawed at passage #8 and propagated for 2 passages prior to seeding for viral infection. At the time of seeding, a cell suspension was made in pre-warmed complete medium and digitally counted using the Scepter 3.0 Handheld Cell Counter and 60 µm Counter Sensors (Millipore, Cat # PHCC360KIT). Cells were seeded for viral infection at passage #11 into replicate tissue culture flasks and multi-well plates using $5x10^6$ cells per T175 flask or $5x10^5$ cells per well in 12-well plates. Seeded flasks and plates were incubated for 24 hrs at 37°C, yielding $\sim1x10^7$ cells per flask and $\sim1x10^6$ cells per well at the time of infection.

Viral infection and cell harvesting

Modified Vaccinia Ankara virus, strain VACV-MVA, was obtained from the ATCC (Cat # VR-1508). VACV-MVA viral stock was diluted in infection medium consisting of DMEM

supplemented with 2% (v/v) FBS. Vero cells were infected at 0 hrs using 0.4 MOI in a minimal volume of inoculum (3 mL per flask) for 1 hr at 37°C; 17 mL infection medium was added prior to incubation for 12-24 hrs at 37°C. In parallel to the flasks, 12-well plates for ICC staining and cell imaging were also infected in parallel to the flasks at 0 hr using 0.4 MOI in 300 µL per well; then 700 µL infection medium was added prior to incubation for 12-24 hrs at 37°C.

Cells were harvested in two biological replicates (A and B) per condition (mock-infected and MVA-infected) at each timepoint (12 hrs, 18 hrs, and 24 hrs post-infection). At each timepoint, cells were washed using 1X phosphate-buffered saline (PBS, Gibco, Cat # 10010049), detached for 15 min at 37°C using 1X TrypLE Express Enzyme (Gibco, Cat # 12604021), and resuspended in pre-warmed complete medium before cell counting with the Scepter 3.0 Handheld Cell Counter and 60 µm Counter Sensors. At each timepoint, mock-infected and MVA-infected cell suspensions were divided and processed for Hi-C, ATAC-seq and RNA-seq techniques (described in section: Library preparation: HiC, ATAC-seq, and RNA-seq).

Immunocytochemistry (ICC) staining and cell imaging

At 12, 18, and 24 hrs post-infection, a 12-well plate of Vero cells was washed with 1X Hank's Balanced Salt Solution (HBSS) containing calcium and magnesium (Gibco, Cat # 14025092), then cross-linked for 15 min using a 4% (w/v) solution of methanol-free formaldehyde in PBS (Invitrogen, Cat # FB002). Cross-linked cells were washed twice in 1X PBS, then stored in 1X HBSS at 4°C for 1-2 days. Subsequent ICC staining steps were performed at room temperature. Cells were quenched of residual formaldehyde for 15 min using 0.01% (v/v) Tween-20 in PBS containing 125 mM glycine, then washed twice in 1X PBS. Cells were permeabilized for 15 min using 0.2% TritonX-100 in PBS, then washed twice in 1X PBS. Cells

31

were blocked for 1-2 hrs using 2% (v/v) Bovine Serum Albumin (BSA, Thermo Scientific, Cat #

37525) + 5% FBS in HBSS, then washed twice in 1X PBS. Five drops of Staining Enhancer

solution (Thermo Scientific, Cat # I36933) was added to cells for 30 min, then washed twice in

1X PBS. Cells were stained with a mouse anti-vaccinia virus monoclonal antibody (Santa Cruz,

Cat # sc-58210), diluted to 1 μg/mL in 0.02% BSA in HBSS, for 24 hrs at 4°C, then washed

twice in 0.01% Tween-20 in PBS. Cells were stained with a rabbit anti-mouse IgG-FITC

secondary antibody (Santa Cruz, Cat # sc-358916), diluted to 1 μg/mL in 0.02% BSA in HBSS,

for 1 hr, washed twice in 0.01% Tween-20 in PBS, then washed twice in 1X PBS to remove

residual Tween. Cells were stored in SlowFade Diamond Anti-Fade Mountant (Invitrogen, Cat #

S36972) at 4°C. Cell imaging was performed within 2 days of ICC staining using an Invitrogen

EVOS FLoid imaging system.

Library preparation: Hi-C, ATAC-seq, and RNA-seq

For Hi-C, $1x10^6$ cells per aliquot were processed in triplicate on the day of cell harvest. Cells

were resuspended in 1X PBS and cross-linked for exactly 10 min using 37% (w/v) formaldehyde

containing 10-15% (v/v) methanol (Sigma, Cat # 252549) at a final concentration of 2%

(v/v). Formaldehyde was quenched for 5 min using glycine at a final concentration of 0.2

M. Cross-linked cells were incubated on ice for an additional 15 min, then pelleted by

centrifugation and washed in 1X PBS. Washed cells were pelleted by centrifugation, and the dry

pellets were snap-frozen on dry ice for long-term storage. These cross-linked cells were

subsequently processed for Hi-C analysis per the Arima Hi-C Kit manual (Cat # A160134) using

the reagents provided in the kit. In summary, cross-linked cells were lysed and enzymatically

treated to create proximally-ligated DNA, which was then purified via AMPure XP Reagent

(Beckman, Cat # A63880). This proximally-ligated DNA was used to create DNA libraries for

32

sequencing per the Arima Hi-C Library Preparation user guide (Cat # A160141). DNA libraries were then sequenced via the Illumina NextSeq 2000 sequencer platform at the Los Alamos National Laboratory Genome Center.

For ATAC-seq (Assay for Transposase-Accessible Chromatin by sequencing), $2x10^5$ cells were cryopreserved in triplicate 1 mL aliquots using freeze medium on the day of cell harvest. Aliquots were subsequently thawed into complete medium, counted to achieve 80,000 cells per sample, and processed for ATAC-seq analysis per the Active Motif ATAC-seq Kit manual (Cat # 53150) using the reagents provided in the kit. In summary, cells were washed in ice-cold 1X PBS, then lysed in ice-cold 'ATAC Lysis Buffer', and centrifuged for 10 min at 4°C. Tagmented DNA was prepared from the cell lysates and purified via DNA purification beads. DNA libraries were prepared from tagmented DNA amplified via PCR using Q5 Polymerase and i7 and i5 Indexed Primers under the following thermocycling conditions: 72 °C for 5 min, 98 °C for 30 sec, 10 cycles of: 98 °C for 10 sec, 63 °C for 30 sec, 72 °C for 1 min, then 10 °C hold. PCR-amplified libraries were cleaned using magnetic DNA purification beads (provided in kit). Additional bead-based size selection was performed using AmpureXP beads (Beckman, Cat # A63880) following manufacturer's protocol to remove fragments below 150bp and above 600bp. DNA libraries were then quantified using a high sensitivity Qubit dsDNA Quantification Assay kit (Thermo Scientific, Cat# Q32851), and KAPA Library Quantification Kit (Roche, Cat # KR0405), and sequenced via the Illumina NextSeq 2000 sequencer platform in paired end mode (PE151).

For RNA-seq, $2x10^5$ cells were cryopreserved in duplicate 1 mL aliquots using freeze medium on the day of cell harvest. Aliquots were subsequently thawed in a 37°C water-bath into ice-cold complete medium, then pelleted by centrifugation, washed in ice-cold HBSS, and

resuspended in 350 µL ice-cold RULT Buffer (Qiagen, Cat # 73934) containing 1% beta-mercaptoethanol (Sigma, Cat #63689). RNA was extracted from these cell lysates using the RNeasy UCP Micro Kit protocol and reagents (Qiagen, Cat #73934). DNA was digested via on-column DNase I treatment for 15 min at room temperature. RNA extracts were eluted in 20 µL RNase-free water. The concentration of total RNA was determined using the Qubit RNA HS Assay Kit (ThermoFisher Scientific, Cat # Q32855) and quality was measured using the Bioanalyzer RNA 6000 Assay (Agilent, Cat. # 5067-1511). Ribosomal RNA was depleted from 500 ng total RNA using the Illumina Ribo-Zero Plus rRNA Depletion Kit (Illumina, Cat. # 20037135). RNA was converted to cDNA, and adapters and indexes were added onto the ends of the fragments to generate strand specific Illumina libraries. Libraries were eluted in DNA Elution Buffer (Zymo Research, Cat #D3004-4-10). The concentration and average size of the libraries was verified using the TapeStation DNA 5000 Assay Reagents and Screen Tapes (Agilent, Cat. # 5067-5589 and 5067-5588). Libraries were normalized and pooled based on the TapeStation concentrations. An accurate library quantification of the pool was determined using the Illumina/Universal Kit (KAPA Biosystems, KK4824). The pool was sequenced on an entire lane of a NextSeq 500 HO flow cell to generate paired-end 151 bp reads using the NextSeq 500 High Output Kit v2.5 (300 cycles) (Illumina, Cat # 20024908).

ATAC-seq analysis

Raw reads from ATAC-seq assays were trimmed and filtered with fastp[56] to remove Nextera adaptors and reads with repetitive sequences. Reads were also filtered to remove bases with low quality scores (q < 15). These processed reads were aligned using bwa[57] to the *Chlorocebus sabaeus* reference genome (GCF_000409795.2_Chlorocebus_sabeus_1.1_genomic.fasta) downloaded from the NCBI genome browser. Prior to alignment, the reference fasta file was

34

modified to include the sequence of the Vaccina virus (Accession No. U94848, downloaded from GenBank). After alignment, duplicate sequenced pairs and mitochondria read pairs were marked and removed from analysis via samblaster[58]. After filtering, sample alignments were analyzed to identify loci displaying significant enrichment of paired-end reads (peaks) using MACS2[59]. Input alignments were further filtered using samtools[60] with the following samtools flags: -F 4 -F 256 -F 512 -F 1024 -F 2048 -q 30 and then passed to MACS2. MAC2 was called in BAMPE mode, with a min-length and max-gap setting of 100 and 150 bp, respectively, in addition to the calling summits option. Replicate correlation analysis was performed with deepTools[61] –plotCorrelation function using MACS2 output bigwig files excluding Vaccinia virus contig as input.

For all further analysis, entire list of MACS2 called peaks except the ones in Vaccinia virus contig were used. List of open chromatin regions (OCRs), defined as regions that are accessible in at least one of sample studied, were created by taking a union of all MACS2 called peaks followed by merging them using BEDtools[62] merge function. To visualize global changes in accessibility, deepTools –computematrix function with –binsize 10kb was used to calculate read coverage around all OCR midpoints. Further, plotHeatmap and plotProfile functions were used for visualization. Differential accessibility analysis across all OCRs was performed using the DESeq2[63] matrix design (~time + treatment).

Hi-C contact map generation and analysis

Hi-C paired fastq.gz files from paired-end Illumina sequencing were analyzed using the Juicer Hi-C pipeline[64]. Samples were aligned to the modified *C. sabaeus* reference genome (described above). Restriction sites associated with the Arima protocol (used to construct Hi-C libraries)

were identified using Juicer. Hi-C files (.hic) were constructed using default settings within Juicer, and for subsequent analysis, the mapping quality threshold of 'q = 30' was used. The number of valid Hi-C contacts across samples, after alignment and filtering, ranged from approximately 256 to 534 million (Supplementary Table 1). The spectral statistic[65] was utilized to measure reproducibility between Hi-C replicates (Figure S1B). The Hi-C replicates demonstrated broad biological reproducibility as demonstrated by this statistic (median scores > 0.75). Replicate Hi-C maps were merged using the "mega" function in Juicer. The number of Hi-C contacts of these merged Hi-C maps ranged from approximately 687 to 854 million.

Distance decay and PCA

The calculation of decay frequencies in Hi-C contacts across genomic distances and principal component analysis (PCA) were conducted in Python using the FAN-C package[66]. Decay profiles were calculated for each chromosome and combined for visualization and analysis. For principal component analysis, a resolution of 50 kb was used to construct genomic bins and the bins with zero Hi-C contacts across all the replicates were removed (-z flag in FAN-C).

Loop identification

Chromatin loops were identified across the replicate and merged Hi-C maps in Juicer via the HiCCUPs[64] algorithm with default settings at resolutions of 5, 10, and 25 kb. To identify differential loops, Hi-C contact counts were gathered for each replicate across the merged list of loops. These counts were fed into DESeq2[63] to identify the following: 1) differential loops between mock-infected control and MVA-infected libraries, 2) loops with differential Hi-C counts across time, and 3) differential loops between mock-infected and MVA-infected libraries.

All analyses were performed across the time course (12, 18, and 24 hpi) and separately within a given timepoint.

A and B compartment assignment and exchange estimation

For each replicate Hi-C map, the initial A and B compartment scores across chromosomes (at a resolution of 100 kb) were estimated using the eigenvector function in Juicer. These results were combined with the replicate ATAC-seq profiles to correctly label active (A) and inactive (B) compartments. Using the convention that the active A compartment should have more ATAC-seq peaks, the number of ATAC-seq peaks in every 100 kb were correlated with the associated A and B scores. For each chromosome, compartment scores were reoriented (i.e., multiplying scores per chromosome by negative one) if the 95th percentile of counts of ATAC-seq peaks were labeled as compartment B in a majority (four out of the six) of ATAC-seq libraries. To identify compartment exchange due to MVA infection, we selected 100 kb bins assigned to the same compartments in both Hi-C biological replicates. Then, the replicate average compartment score was determined for each bin considering only the bins in agreement between replicates. The changes from positive (in control) to negative (in infected) bin score value was assigned as A to B exchange; similarly, changes from negative (in control) to positive (in infected) bin score value were assigned as B to A exchange. Bins that remained positive or negative were assigned as 'no exchange' A or B compartments, respectively.

Inter-chromosome interaction analysis

An inter-chromosomal interaction score was calculated for each pair of chromosomes within each Hi-C experimental condition. As described in detail in Duan et al. (2010), this score is a ratio calculated by dividing the number of Hi-C contacts between a pair of chromosomes by the

expected frequency of inter-chromosomal Hi-C contacts involving those chromosomes. After calculating this ratio for each Hi-C map, the difference (in $\log_2$ scale) in inter-chromosomal score was calculated across the infection time course (12, 18, and 24 hpi) between each pair of mock-infected control and MVA-infected maps.

Differential interaction analysis

multiHiCcompare[36] was used to perform joint normalization (using fastlo method) of the Hi-C matrix across comparative datasets and to identify differential contacts (using hic_exactTest). For input, the Hi-C contact matrix converted to a four-column bed file (chromosome, region1, region2, IF) in 10 kb contiguous windows across the genome was used.

Pathway enrichment prediction based on accessibility

Lists of genes whose promoter (TSS $\pm$ 500 bp) overlapped with a strong differential peak (fold change >1 and base mean >100) were prepared. This type of conservative list, predicting altered gene transcription due to infection, was generated for both MVA infection accessible and control accessible genes for each time point. KOBAS[41] software was used to identify the statistically significant KEGG pathways enriched in each list using the entire list of genes in *Chlorocebus sabaeus* as background.

RNA-seq analysis

The *Chlorocebus sabaeus* reference genome was retrieved from NCBI (GCF_000409795.2) and appended with the vaccinia virus (VACV) genome, retrieved from GenBank (U94848.1). Genome annotations for *C. sabaeus* were downloaded from NCBI and the GenBank flat file annotations for VACV were retrieved using the bio[67] package command (bio fetch U94848.1 --

format gff) and converted to gtf format using the AGAT[68] package function (agat_convert_sp_gff2gtf.pl). The *C. sabaeus* annotation was appended with the VACV annotation for subsequent analysis. Raw read quality control was performed using FastQC[69] and SeqMonk[70] to verify library quality and degree of PCR duplication. All metrics were normal. A splice junction aware reference index was generated for the combined genome and annotation files using STAR[71] with the arguments (--runMode genomeGenerate --sjdbOverhang 149). Pre-alignment read trimming was not performed in favor of soft-clipping/quality filtering by the STAR aligner. Reads were aligned to the combined genome using STAR. Duplicate reads were marked but retained in downstream analysis per DESeq2 usage guidelines. Sense strand reads aligning to gene features were counted using htseq-count[72] with the arguments (--stranded=reverse). Antisense strand reads aligning to gene features were counted using htseq-count with the arguments (--stranded=yes). Differential expression analysis was performed using the DESeq2[63] matrix design (~time + treatment). For visualization, replicate alignment files were merged using samtools. bigWigs were then generated using the bamCoverage function from the deepTools[61] package with the arguments (--effectiveGenomeSize 2836634000 --normalizeUsing BPM –filterRNAstrand {forward/reverse/NULL} --binSize 1).

Gene enrichment analysis

Up and down regulated genes, with adjusted *p*-value <0.05 and $\log_2$ fold change greater than 2 or less than 2 respectively were used. This type of conservative list, representing altered gene transcription due to infection, was generated for both MVA infection and control datasets for each time point. KOBAS[41] software was used to identify the statistically significant KEGG pathways enriched in each list using the entire list of genes in *Chlorocebus sabaeus* as background.

Accessibility, gene expression, chromatin loop integration analysis

Accessibility across differential loop anchor intervals and around TSS of differential expressed genes were calculated using deepTools[61] –computematrix function with –binsize 10kb, skipping bins with zero read coverage. For estimating average treatment by time accessibility across all control-biased (p_{adj}<0.05 and log$_2$foldchange <0) and infection-biased (p_{adj}<0.05 and log$_2$foldchange >0) loop anchors --scale-regions mode was used. Whereas, --referencepoint mode was used for to estimate accessibility around TSS of upregulated genes (p_{adj}<0.05 and log$_2$foldchange >2) and downregulated genes (p_{adj}<0.05 and log$_2$foldchange <2) at 24 hpi. plotHeatmap and plotProfile functions were used for visualization.

To estimate the fraction of differential OCRs and genes that overlaps with gained/restructured loop domains due to infection, considered all infection-biased loops (p_{adj}<0.05 and log$_2$foldchange >0) and assigned the region between left boundary of left anchor interval and right boundary of right anchor interval as gained/restructured loop domain. Further, BEDtools[62] intersect function was used to estimate the number of decreased accessibility OCRs (p_{adj}<0.05, baseMean > 100, log$_2$foldchange < -1) and increased accessibility OCRs (p_{adj}<0.05, baseMean > 100, log$_2$foldchange > 1) overlapping with gained/restructured loop domains. Similarly, the overlaps of upregulated genes (p_{adj}<0.05, log$_2$foldchange > 2) and downregulated genes (p_{adj}<0.05, log$_2$foldchange <-2) were also calculated. To obtain a null distribution, differential OCRs and genes were shuffled 10,000 times across the entire *Chlorocebus sabaeus* genome using BEDtools shuffle function and the overlap with gained/restructured loop domain with each iteration was calculated. 'empPvals' function from R package 'qvalue' was used to estimate the associated empirical *p*-value.

Data Availability

All data are available under Bioproject #XXXXXX and associated SRA #XXXXXX. Schematics are created using Biorender.com.

Author Contributions

CRS and VV designed the viral experiments and chromatin architecture, accessibility, and expression assays and analyses. SA performed the VACV viral infections and microscopy. SA and VV performed cell harvest and sample collection. SA performed the RNA extractions and the LANL Genome Center prepared the RNA-seq libraries. VV performed sample processing and preparation of ATAC-seq and Hi-C libraries. All sequencing was conducted at the LANL Genome Center. CR aligned and processed the Hi-C sequencing data. CR and VV analyzed the Hi-C data. CR and VV developed the pipeline for ATAC-seq analysis. VV analyzed the ATAC-seq data, including visualization and interpretation. ES processed the RNA-seq data. VV interpreted and integrated all datasets (Hi-C, ATAC-seq, and RNA-seq). VV and CR generated the figures and tables. VV and CRS performed analyses and biological interpretation of all datasets and wrote the manuscript. Funding was provided by CRS, SS, and KYS.

Acknowledgements

References

1   Rao, S. S. *et al.* A 3d map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).

2   Misteli, T. Beyond the sequence: Cellular organization of genome function. *Cell* **128**, 787-800 (2007).

3   Keizer, V. I. P. *et al.* Live-cell micromanipulation of a genomic locus reveals interphase chromatin mechanics. *Science* **377**, 489-495,(2022).

4   Meijering, A. E. C. *et al.* Nonlinear mechanics of human mitotic chromosomes. *Nature* **605**, 545-550 (2022).

5   Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13-23 (2017).

6   Parmar, J. J., Woringer, M. & Zimmer, C. How the genome folds: The biophysics of four-dimensional chromatin organization. *Annu Rev Biophys* **48**, 231-253 (2019).

7   Lappala, A. *et al.* Four-dimensional chromosome reconstruction elucidates the spatiotemporal reorganization of the mammalian X chromosome. *Proc Natl Acad Sci U S A* **118**, e2107092118 (2021).

8   Heinz, S. *et al.* Transcription elongation can affect genome 3D structure. *Cell* **174**, 1522-1536 (2018).

9   Rowley, M. J. *et al.* Evolutionarily conserved principles predict 3D chromatin organization. *Mol Cell* **67**, 837-852 (2017).

10   Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418-422 (2016).

11   Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331-336 (2015).

12   Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948-953 (2013).

13   Wang, R. *et al.* SARS-CoV-2 restructures host chromatin architecture. *Nat Microbiol* **8**, 679-694 (2023).

14   Wang, T. *et al.* Chemical-induced phase transition and global conformational reorganization of chromatin. *Nat Commun* **14**, 5556 (2023).

15   Song, Z. T., Liu, J. X. & Han, J. J. Chromatin remodeling factors regulate environmental stress responses in plants. *J Integr Plant Biol* **63**, 438-450 (2021).

16   Silmon de Monerri, N. C. & Kim, K. Pathogens hijack the epigenome: A new twist on host-pathogen interactions. *Am J Pathol* **184**, 897-911 (2014).

17   Lieberman, P. M. Chromatin organization and virus gene expression. *J Cell Physiol* **216**, 295-302 (2008).

18   Majumder, K. & Morales, A. J. Utilization of host cell chromosome conformation by viral pathogens: Knowing when to hold and when to fold. *Front Immunol* **12**, 633762 (2021).

19   Menachery, V. D. *et al.* MERS-CoV and H5N1 influenza virus antagonize antigen presentation by altering the epigenetic landscape. *P Natl Acad Sci USA* **115**, E1012-E1021 (2018).

20   Isidro, J. *et al.* Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. *Nat Med* **28**, 1569-1572 (2022).

21   Greseth, M. D. & Traktman, P. The life cycle of the vaccinia virus genome. *Annu Rev Virol* **9**, 239-259, (2022).

22   Hruby, D. E., Guarino, L. A. & Kates, J. R. Vaccinia virus replication. I. Requirement for the host-cell nucleus. *J Virol* **29**, 705-715 (1979).

23   Moyer, R. W. The role of the host cell nucleus in vaccinia virus morphogenesis. *Virus Res* **8**, 173-191 (1987).

24   Konishi, K. *et al.* Whole-genome sequencing of Vero E6 (VERO C1008) and comparative analysis of four Vero cell sublines. *Front Genet* **13**, 801382 (2022).

25   Okeke, M. I. *et al.* Hazard characterization of modified vaccinia virus ankara vector: What are the knowledge gaps? *Viruses* **9**, 318 (2017).

26   Kiesslich, S. & Kamen, A. A. Vero cell upstream bioprocess development for the production of viral vectors and vaccines. *Biotechnol Adv* **44**, 107608 (2020).

27   Volz, A. & Sutter, G. Modified vaccinia virus ankara: History, value in basic research, and current perspectives for vaccine development. *Adv Virus Res* **97**, 187-243 (2017).

28   Rerks-Ngarm, S. *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* **361**, 2209-2220 (2009).

29   Perez, P. *et al.* A vaccine based on a modified vaccinia virus ankara vector expressing zika virus structural proteins controls zika virus replication in mice. *Sci Rep* **8**, 17385 (2018).

30   Boudewijns, R. *et al.* MVA-CoV2-S vaccine candidate neutralizes distinct variants of concern and protects against SARS-CoV-2 infection in Hamsters. *Front Immunol* **13**, 845969 (2022).

31   Meseda, C. A. *et al.* MVA vector expression of SARS-CoV-2 spike protein and protection of adult Syrian hamsters against SARS-CoV-2 challenge. *NPJ Vaccines* **6**, 145 (2021).

32   Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293 (2009).

33   Nikitina, T., Norouzi, D., Grigoryev, S. A. & Zhurkin, V. B. DNA topology in chromatin is defined by nucleosome spacing. *Sci Adv* **3**, e1700957 (2017).

34   Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486-490 (2015).

35   Oluwadare, O., Highsmith, M. & Cheng, J. An overview of methods for reconstructing 3-D chromosome and genome structures from Hi-C data. *Biol Proced Online* **21**, 7 (2019).

36   Stansfield, J. C., Cresswell, K. G. & Dozmorov, M. G. multiHiCcompare: joint normalization and comparative analysis of complex Hi-C experiments. *Bioinformatics* **35**, 2916-2923 (2019).

37   Brahmachari, S. & Marko, J. F. Chromosome disentanglement driven via optimal compaction of loop-extruded brush structures. *Proc Natl Acad Sci U S A* **116**, 24956-24965 (2019).

38   Teves, S. S., Weber, C. M. & Henikoff, S. Transcribing through the nucleosome. *Trends Biochem Sci* **39**, 577-586 (2014).

39   Nocetti, N. & Whitehouse, I. Nucleosome repositioning underlies dynamic gene expression. *Genes Dev* **30**, 660-672 (2016).

40   Stewart-Morgan, K. R., Reveron-Gomez, N. & Groth, A. Transcription restart establishes chromatin accessibility after DNA replication. *Mol Cell* **75**, 408-414 (2019).

41   Bu, D. *et al.* KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res* **49**, W317-W325 (2021).

42   Kiani, K., Sanford, E. M., Goyal, Y. & Raj, A. Changes in chromatin accessibility are not concordant with transcriptional changes for single-factor perturbations. *Mol Syst Biol* **18**, e10979 (2022).

43   Mirzaei, H. & Faghihloo, E. Viruses as key modulators of the TGF-β pathway; a double-edged sword involved in cancer. *Rev Med Virol* **28**, e1967 (2018).

44   Wang, Z. *et al.* The Hippo pathway and viral infections. *Front Microbiol* **10**, 3033 (2019).

45   Dunn, E. F. & Connor, J. H. HijAkt: The PI3K/Akt pathway in virus replication and pathogenesis. *Prog Mol Biol Transl Sci* **106**, 223-250 (2012).

46   Sumbria, D., Berber, E., Mathayan, M. & Rouse, B. T. Virus infections and host metabolism-can we manage the interactions? *Frontiers in Immunology* **11**, 594963 (2021).

47   Riedmann, C. *et al.* Inorganic Arsenic-induced cellular transformation is coupled with genome wide changes in chromatin structure, transcriptome and splicing patterns. *Bmc Genomics* **16**, 212 (2015).

48   Scott, A. K., Rafuse, M. & Neu, C. P. Mechanically induced alterations in chromatin architecture guide the balance between cell plasticity and mechanical memory. *Front Cell Dev Biol* **11**, 1084759 (2023).

49   Anania, C. & Lupianez, D. G. Order and disorder: Abnormal 3D chromatin organization in human disease. *Brief Funct Genomics* **19**, 128-138 (2020).

50   Liang, W. Z. *et al.* When 3D genome technology meets viral infection, including SARS-CoV-2. *J Med Virol* **94**, 5627-5639 (2022).

51   Smith, G. L., Talbot-Cooper, C. & Lu, Y. X. How does vaccinia virus interfere with interferon? *Advances in Virus Research* **100**, 355-378 (2018).

52   Schmitt, A. D. *et al.* A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep* **17**, 2042-2059 (2016).

53   Ma, F. *et al.* Contribution of carbohydrate-related metabolism in Herpesvirus infections. *Curr Res Microb Sci* **4**, 100192 (2023).

54   Turjya, R. R., Khan, M. A. A. & Islam, A. B. M. M. K. Perversely expressed long noncoding rnas can alter host response and viral proliferation in SARS-CoV-2 infection. *Future Virol* **15**, 577-593 (2020).

55   Misteli, T. The self-organizing genome: Principles of genome architecture and function. *Cell* **183**, 28-45 (2020).

56   Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).

57   Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

58   Faust, G. G. & Hall, I. M. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* **30**, 2503-2505 (2014).

59   Zhang, Y. *et al.* Model-based Analysis of Chip-Seq (MACS). *Genome Biol* **9**, R137 (2008).

60   Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

61   Ramirez, F. *et al.* deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).

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

63   Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

64   Durand, N. C. *et al.* Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst* **3**, 95-98 (2016).

65   Yan, K. K., Yardimci, G. G., Yan, C., Noble, W. S. & Gerstein, M. HiC-spector: A matrix library for spectral and reproducibility analysis of Hi-C contact maps. *Bioinformatics* **33**, 2199-2201 (2017).

66   Kruse, K., Hug, C. B. & Vaquerizas, J. M. FAN-C: A feature-rich framework for the analysis and visualisation of chromosome conformation capture data. *Genome Biol* **21**, 303 (2020).

67   Parnell, L. D. *et al.* BioStar: An online question & answer resource for the bioinformatics community. *PLoS Comput Biol* **7**, e1002216 (2011).

68   Dainat, J. *et al.* Nbisweden/agat: Agat-v1.2.0 (v1.2.0), Zenodo. doi:10.5281/zenodo.8178877 (2023).

69   Andrews, S. FastQC: A quality control tool for high throughput sequence data [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).

70   Andrews, S. SeqMonk. https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/ (2007).

71   Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

72   Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E. & Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38**, 2943-2945 (2022).