Article

SOURCE DATA | TRANSPARENT PROCESS | OPEN ACCESS

THE EMBO JOURNAL

Changes in adenoviral chromatin organization precede early gene activation upon infection

Uwe Schwartz[1], Tetsuro Komatsu[2,3], Claudia Huber[1], Floriane Lagadec[3,4], Conradin Baumgartl[1], Elisabeth Silberhorn[1], Margit Nuetzel[5], Fabienne Rayne[3], Eugenia Basyuk[3], Edouard Bertrand[6], Michael Rehli[5,7,8], Harald Wodrich[3,*] & Gernot Laengst[1,**]

Abstract

Within the virion, adenovirus DNA associates with the virus-encoded, protamine-like structural protein pVII. Whether this association is organized, and how genome packaging changes during infection and subsequent transcriptional activation is currently unclear. Here, we combined RNA-seq, MNase-seq, ChIP-seq, and single genome imaging during early adenovirus infection to unveil the structure- and time-resolved dynamics of viral chromatin changes as well as their correlation with gene transcription. Our MNase mapping data indicates that the adenoviral genome is arranged in precisely positioned nucleoprotein particles with nucleosome-like characteristics, that we term adenosomes. We identified 238 adenosomes that are positioned by a DNA sequence code and protect about 60–70 bp of DNA. The incoming adenoviral genome is more accessible at early gene loci that undergo additional chromatin de-condensation upon infection. Histone H3.3 containing nucleosomes specifically replaces pVII at distinct genomic sites and at the transcription start sites of early genes. Acetylation of H3.3 is predominant at the transcription start sites and precedes transcriptional activation. Based on our results, we propose a central role for the viral pVII nucleoprotein architecture, which is required for the dynamic structural changes during early infection, including the regulation of nucleosome assembly prior to transcription initiation. Our study thus may aid the rational development of recombinant adenoviral vectors exhibiting sustained expression in gene therapy.

Keywords adenovirus; chromatin remodeling; early infection; pVII; transcription

Subject Categories Chromatin, Transcription & Genomics; Microbiology, Virology & Host Pathogen Interaction

DOI 10.15252/embj.2023114162 | Received 31 March 2023 | Revised 10 July 2023 | Accepted 4 August 2023 | Published online 29 August 2023

The EMBO Journal (2023) 42: e114162

Introduction

Most DNA viruses translocate their genomes into the nucleus of infected target cells to initiate viral gene expression and genome replication. During cytoplasmic trafficking, genomes are highly compacted inside the viral capsid to protect them from nucleic acid sensing molecules. It is generally assumed that genomes are decondensed upon entering the nucleus and thereby activated for viral gene expression. In the nucleus depending on the virus and cell context viral latency is established or replication initiated. Replicated viral genomes would then need to be condensed for packaging into progeny. Alternatively, nuclear viral genomes could undergo different parallel fates, i.e. only one part serves as transcription unit, whereas others are replicated and/or packaged (Brison *et al*, 1979). While these remarkable and reversible genome functions are highly specialized and efficient, for many viruses little is known about the structural organization of the genome, its dynamic changes or the regulatory cues initiating structural transitions during virus entry.

Adenoviruses (Ad) are DNA viruses containing a linear double stranded DNA genome of 30–40 kb that is incorporated into a ~90 nm non-enveloped icosahedral capsid shell. Ad enter cells by receptor-mediated endocytosis. Stepwise entry cues liberate the internal capsid protein VI and permit the virus to escape the endosomal compartment (Greber *et al*, 1993; Wickham *et al*, 1993; Wiethoff *et al*, 2005). Cytoplasmic capsids are transported to the nuclear pore complex (NPC) where they dock, disassemble and release the viral genome for nuclear import (Strunze *et al*, 2011; Cassany *et al*, 2015). Within hours after cell entry, nuclear genomes initiate gene expression from

1 Biochemie Zentrum Regensburg, University of Regensburg, Regensburg, Germany
2 Laboratory of Epigenetics and Metabolism, Institute for Molecular and Cellular Regulation, Gunma University, Gunma, Japan
3 CNRS UMR 5234, Microbiologie Fondamentale et Pathogénicité, Université de Bordeaux, Bordeaux, France
4 Department of Molecular Biology, Faculty of Medicine, Göttingen Center of Biosciences (GZMB), Georg-August-University Göttingen, Göttingen, Germany
5 Department of Internal Medicine III, University Hospital Regensburg, Regensburg, Germany
6 CNRS UMR 5355, Institut de Généthique Moléculaire de Montpellier, Montpellier, France
7 Leibniz Institute for Immunotherapy, Regensburg, Germany
8 University Hospital Regensburg, Regensburg, Germany
*Corresponding author. Tel: +33 557 571130; E-mail: harald.wodrich@u-bordeaux.fr
**Corresponding author. Tel: +49 941 9432600; E-mail: gernot.laengst@ur.de

the immediate early promoter. In a simplified view this promoter encodes the E1A gene, encoding a *trans*-activator driving early genes responsible for cell and immune modulation (E1B, E3, E4) and replication (E2). In contrast, activation of late genes (L1–L5), encoding most structural proteins, requires prior genome replication (Flint, 1982; Akusjarvi, 2008).

Packaged Ad genomes are devoid of histones. Instead, they associate with basic core proteins VII (pVII), V (pV) and polypeptide μ (Brown *et al*, 1975; Corden *et al*, 1976). Protein VII is the most tightly bound and most abundant protein with over 500 copies per virion and responsible for packaging the viral genome into a structure appearing as irregular "beads-on-a-string" by electron microscopy (Benevento *et al*, 2014). Protein VII exhibits sequence similarities to protamines and may use its positive charges to neutralize and compact the viral DNA (Anderson *et al*, 1989; Keller *et al*, 2002; Johnson *et al*, 2004). Cellular factors and/or post-transcriptional modification control the association of pVII with viral genomes in the producer cell prior or during genome packaging (Samad *et al*, 2012; Avgousti *et al*, 2016; Inturi *et al*, 2017; Ostapchuk *et al*, 2017; Mun & Punga, 2018; Genoveso *et al*, 2020). A recent study showed that viral genomes can be packaged without pVII, but that its presence is a prerequisite for infection (Ostapchuk *et al*, 2017). The genomic binding sites of pVII as well as the structural organization it imposes on viral DNA, inside the virion, post infection and post nuclear genome delivery, remain elusive (Mirza & Weber, 1982; Giberson *et al*, 2011). Protein VII stays associated with the viral genome imported to the nucleus. In contrast, protein V dissociates from the viral DNA following ubiquitylation prior to genome nuclear import (Puntener *et al*, 2011; Bauer *et al*, 2021). The fate of the μ-peptide is not known.

Cellular TAF-Iβ/SET associates with pVII on nuclear viral genomes and is thought to stimulate transcription and prevent the activation of DNA damage responses (Haruki *et al*, 2003; Giberson *et al*, 2011; Komatsu *et al*, 2015; Avgousti *et al*, 2017). Protein VII was also shown to functionally interact with the viral *trans*-activator E1A, pinpointing to a direct role in transcriptional activation (Johnson *et al*, 2004). Some studies suggested that pVII remains attached throughout the early phase of infection while others argue for a gradual and at least partial removal from the incoming genome (Haruki *et al*, 2003; Johnson *et al*, 2004; Xue *et al*, 2005; Komatsu *et al*, 2011, 2015). Conflicting reports postulated that pVII turnover either requires transcription (Chen *et al*, 2007) or occurs independent of transcription, at least on viral vector genomes (Ross *et al*, 2011). In contrast chromatin immune precipitation (ChIP) analysis proposed that pVII remains associated with viral chromatin for several hours (Chatterjee *et al*, 1986; Haruki *et al*, 2003; Johnson *et al*, 2004; Xue *et al*, 2005; Komatsu *et al*, 2011). In addition, cellular histones may associate with incoming genomes prior to replication (Ross *et al*, 2011; Giberson *et al*, 2018).

Functional chromatin extraction using high and low concentrations of micrococcal nuclease (MNase) to hydrolyze cellular chromatin, combined with high-throughput sequencing, is a way to analyze protein binding to DNA and to reveal dynamic changes in chromatin structure (Mueller *et al*, 2017; Schwartz *et al*, 2018). In this study, we combined time-resolved MNase, ChIP and RNA sequencing as well as single molecule imaging to correlate chromatin structure changes with transcription at very early time points of adenoviral infection. This approach revealed the specific nucleoprotein architecture of the virus, which dynamically changes during early infection. Protein VII positioning is driven by sequence and is specifically organized at early genes, correlating with the early de-condensation of these regions upon infection. Defined pVII nucleoprotein complexes are replaced by nucleosomes containing H3.3 on early gene promoters and few additional genomic sites. pVII remodeling, nucleosome assembly and the specific acetylation of the promoter nucleosomes are preceding transcriptional activation. The time resolved single molecule imaging of transcribing genomes support our findings, identifying the morphological correlate at the single genome level at subcellular resolution. The defined positioning of pVII on the viral genome and its functional involvement in dynamic chromatin rearrangement during infection, suggest a key role in regulating the viral life cycle. We identify a DNA sequence driven adenovirus nucleoprotein architecture, which may aid the design of improved Ad vectors for gene-therapy or vaccination.

Results

Time resolved analysis of the early adenovirus transcriptome by RNA-sequencing

To investigate the early spatio-temporal reorganization of viral chromatin and transcription activation we infected H1299, a p53-deficient non-small cell lung carcinoma cell line, or U2OS, a human osteosarcoma cell line, with a partially E3-deleted but replication competent human adenovirus C5 (HAd-C5dE3). The inoculum was replaced with fresh medium after 30 min, defining time point zero of the experimental setup (Fig 1A). On time point 0, 0.5, 1, 2, and 4 h post infection (hpi) infected cells were subjected to immune fluorescence analysis (IF), RNA-sequencing (RNA-seq), ChIP-seq and functional chromatin extraction combined with high-throughput sequencing of protected DNA fragments (MNase-seq) (Fig 1A). To verify the progression of infection, cells were first analyzed by immunofluorescence (Fig EV1A) for the internal capsid protein VI (pVI), which can be transiently detected during endosomal escape (Martinez *et al*, 2015), while nuclear import is marked by accessibility of the genome bound pVII in the nucleus (Komatsu *et al*, 2015). Protein VI and pVII quantification showed that endosomal escape occurred within the first 30 min (Fig EV1A and B) followed by rapid nuclear delivery of viral genomes starting at 0.5 hpi, with most genomes being imported 1 hpi (Fig EV1A and C). The data are confirming a synchronized infection covering pre- and post-nuclear genome delivery phases.

To investigate viral transcription, we performed mRNA sequencing of two independent infection time-courses yielding on average 26 million reads per sample (Dataset EV1). The reads were mapped to an extended version of the human genome containing the HAd-C5dE3 reference genome and gene annotation. The first mature Ad transcripts were detected at 2 hpi, and strongly increased levels were observed 4 hpi (Fig 1B and C; whole genome profile shown as individual replicates in Appendix Fig S1A and quantification in Dataset EV1). Consistent with the spatial organization of the Ad transcription program (Glenn & Ricciardi, 1988; Crisostomo *et al*, 2019), transcription is initiated at both ends of the Ad genome covering the early genes (Fig 1B). We detected E1A as the first viral transcript (Nevins *et al*, 1979), along with significant levels of E3 at

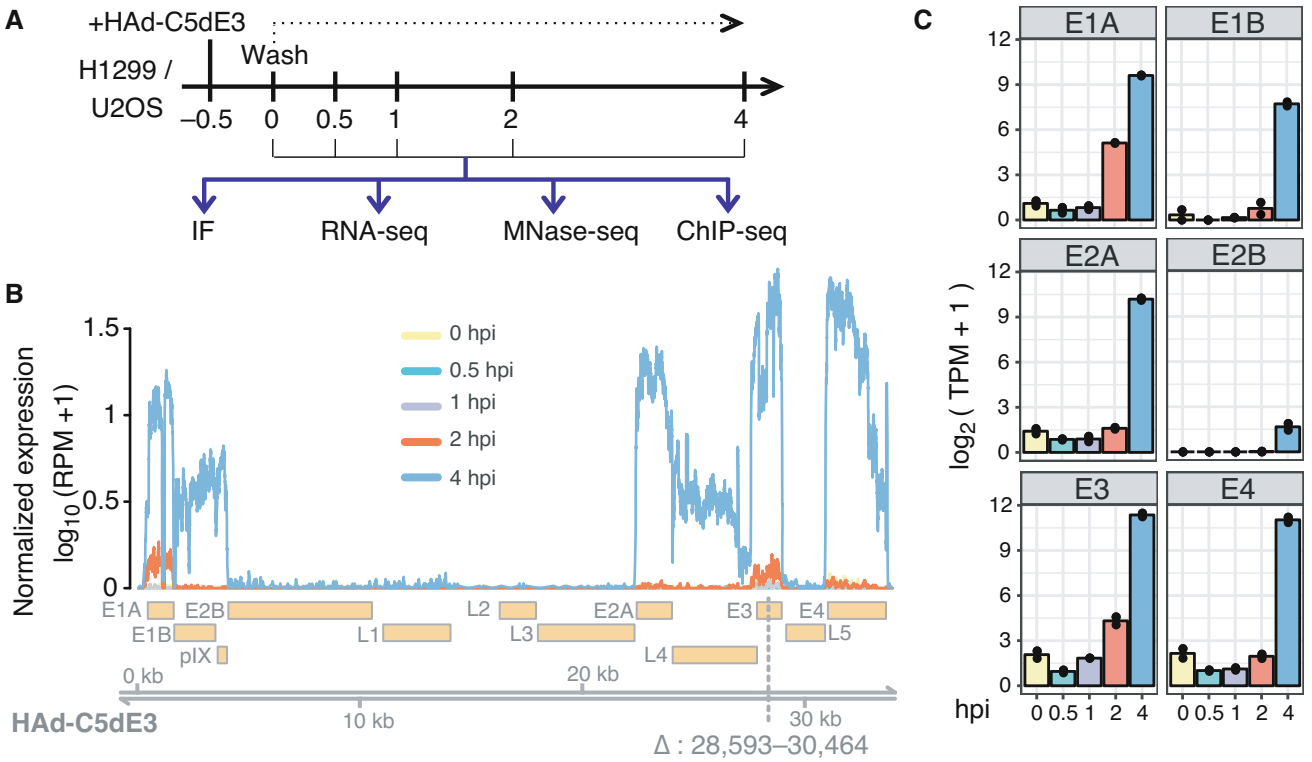

Figure 1. Spatio-temporal analysis of adenovirus transcription using RNA-seq.

A Experimental design. Virus was added (+Ad) and removed (Wash) after 30 min and cell were processed as indicated (hpi = hours post infection) using immunofluorescence (IF), RNA sequencing (RNA-seq), nucleosome sequencing (MNase-seq) and ChIP sequencing (ChIP-seq).

B Profile illustrating Ad genome coverage by RNA-seq reads. Reads were normalized to RPM (reads per million mapped reads) and the average of two replicates is shown. Genomic location and underlying gene annotation of HAd-C5dE3 (NCBI accession: AY339865.1) are indicated at the bottom. The area deleted in the HAd-C5dE3 strain is indicated by the triangle.

C Transcript abundance estimation of early Ad genes. TPMs (Transcripts Per Kilobase Million) for early expressed genes were calculated at each timepoint. The barplots show the average of two replicates and each replicate is indicated by a black dot.

2 hpi (Fig 1B and C). At 4 hpi viral transcription is markedly enhanced with high levels of transcripts originating from all early genes and significant signals arising from the E2A region. Taken together we observed an efficient utilization of the host cell transcription machinery within the first hours of infection, revealing a transcriptional competent viral template.

In situ detection of E1A mRNA identifies transcribing Ad genomes

To directly correlate the progression of infection and the onset of transcription, the import of genome associated pVII and the expression of the E1A gene was monitored simultaneously by immunofluorescence (Fig 2A). We used RNAscope technology for the direct detection of accumulating individual viral transcripts at the single cell level (Pheasant *et al*, 2018). A specific set of probes for the E1A mRNA was designed (ACDBio) and validated on cells infected with either HAd-C5dE3 virus or an E1-deleted GFP expressing vector (Ad5-GFP). At 4 hpi specific E1A transcripts could be readily detected in virus- but not in E1-deleted GFP expressing viral vector infected cells (Fig 2B and Appendix Fig S1B). Protein VII immunofluorescence analysis showed that most genomes were imported before 1 hpi with no significant additional import after this time point (Fig 2C). While the nuclear import of genomes was saturated 1 hpi, RNA-seq and RNAscope revealed that E1A transcripts were detected only at 2 hpi and accumulated at 4 hpi (Figs 2D and 1C). Ad transcription initiation and mRNA synthesis exhibited a significant lag time after nuclear entry, suggesting a requirement for remodeling processes prior to the assembly of the transcription machinery.

E1A transcript accumulation could be related to the continued expression from a stable pool of transcriptionally active genomes or by an increasing number of genomes initiating transcription over time. We quantified the colocalization of nuclear genomes with nuclear E1A transcripts to discriminate between both possibilities. Colocalizations inside the nucleus resulting from association by chance are expected to increase over time when active genome numbers remain constant and E1A nuclear transcript numbers increase over time. However, our results showed that the proportion of E1A positive nuclear genomes tents to rapidly plateau at ~20% despite the strong accumulation of nuclear E1A transcripts (Fig 2E). It is shown that only a relatively small subpopulation of imported genomes contributes to early viral gene expression until 4 hpi, which was also observed in a recent study (Suomalainen *et al*, 2020). These observed ~20% transcribing viral genomes could also be the result of a larger proportion of genomes that undergo

transcriptional bursting. This interpretation would not contradict the observed results.

In vivo detection of E1A mRNA identifies a subpopulation of transcribing Ad genomes

The RNAscope analysis confirmed RNA-seq results and indicated, that active viral genomes undergo a lag phase in the nucleus, probably acquiring structural changes that allow to activate early gene expression. To confirm this observation and further characterize transcribing and non-transcribing genomes, we developed an imaging system for living cells. To mark E1A transcripts in living cells, we inserted 24x MS2 binding sites into the 3′UTR of the E1A gene of a replication competent virus. MS2 binding sites are short stem-loop structures derived from the MS2 bacteriophage and can be detected by binding of multiple copies of fluorophore tagged MS2

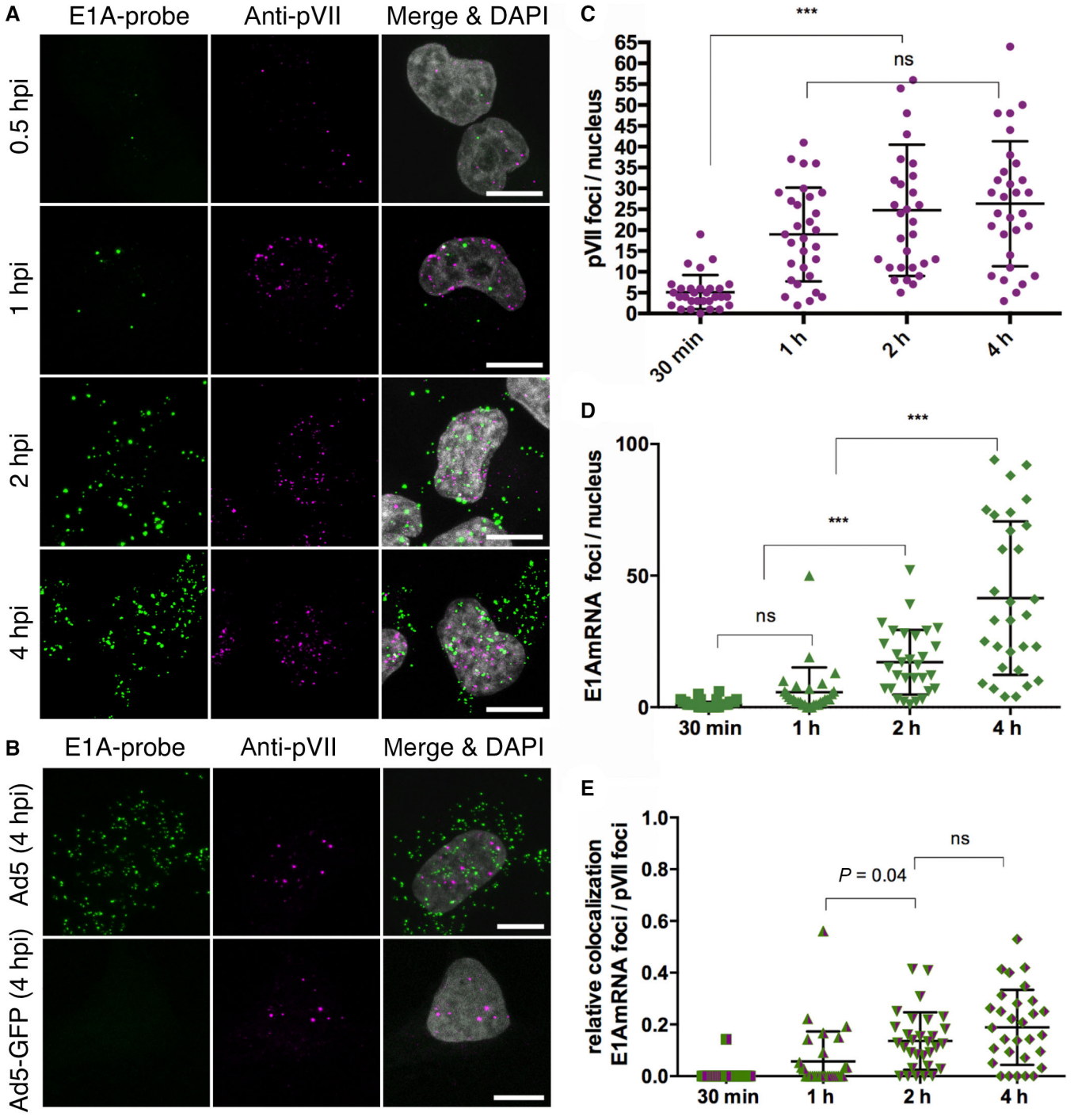

Figure 2.

Figure 2. E1A transcription at the single cell level.

A Representative confocal images of cells infected with HAd-C5dE3 at different time points as indicated. Detection of E1A transcript with specific RNAscope probes (left column, green signal), genomes using anti-pVII antibodies (middle column, magenta signal) and merged with DAPI detecting the nucleus (right column, gray signal). Scale bar is 10 μm.

B Representative confocal images of cells infected with HAd-C5dE3 (Ad5, top row) or an E1-deleted HAd-C5dE3 GFP expressing vector (Ad5-GFP, bottom row). Scale bar is 10 μm.

C Quantification of nuclear pVII foci over time as indicated on the x-axis. The distribution of normalized pVII foci at each time point is shown as scatter plot (Mean ± SD, $n = 30$ cells per time point).

D Graph as in (C) showing the distribution of nuclear E1A transcript foci per cell as scatter plot (Mean ± SD, $n = 30$ cells per time point).

E Scatter plot showing the distribution of nuclear pVII foci (C) positive for E1A transcript foci (D) normalized by total number of nuclear pVII foci (Mean ± SD, $n = 30$ nuclei per time point).

Data information: (C–E) Statistical analysis via one-way ANOVA multi-comparison test, ***$P < 0.001$.
Source data are available online for this figure.

binding protein (i.e. MS2BP-NLS-GFP) forming a fluorescent spot *in situ* (Bertrand *et al*, 1998). Production and amplification of the virus resulted in normal yields with unaltered infectivity or protein composition compared to the parental virus (Fig EV2A). To detect Ad genomes in living cells we used our previously established live cell imaging system based on fluorescent TAF-Iβ, which associates with the pVII of incoming genomes, thereby generating a fluorescent spot *in situ* (Komatsu *et al*, 2015). A stable mCherry-TAF-Iβ expressing cell line was transduced with an expression vector for MS2BP-NLS-GFP to generate double fluorescent stable cell lines. Cells were left either uninfected, or infected with untagged replicative control virus, or infected with E1A-MS2 tagged replicative virus and at ~4 hpi analyzed by live cell imaging at high temporal resolution (1 frame/s) (Movies EV1–EV4). To display the results in a single image we superimposed individual movie frames (60 s t-projections, Fig 3B). In non-infected control cells both, mCherry-TAF-Iβ and MS2BP-NLS-GFP localized to the nucleoplasm, with some unspecific enrichment of the MS2 signal in nucleoli (Fig 3A top row, taken from Movie EV1). When cells were infected with non-MS2 containing control virus, genomes formed mCherry-TAF-Iβ fluorescent spots in the nucleoplasm with limited confined motion and some enrichment in the nuclear periphery as previously described (Komatsu *et al*, 2015), while the MS2 signal remained unchanged compared to non-infected cells (Fig 3A 2nd row taken from Movie EV2). In contrast, cells infected with virus expressing MS2-repeat-tagged E1A mRNA formed mCherry-TAF-Iβ spots indicating viral genomes but also rapidly moving GFP spots visible inside the nucleoplasm and in the cytoplasm (Movies EV3 and EV4). GFP spots accumulated in most cells around 3–5 hpi in agreement with the RNAscope experiment. Upon merging the mCherry-TAF-Iβ channel (viral genomes) with the MS2BP-NLS-GFP channel (viral E1AmRNA) using t-projections, some genomes were clearly double positive suggesting transcribing genomes (Fig 3A 3rd and 4th row, taken from Movies EV3 and EV4). Similar observations were made when we used a different MS2BP construct (MS2BP-NES-GFP) predominantly localizing to the cytosol or when we lowered the input MOI resulting in fewer (< 10) genome copies per imaging field (Fig EV2B). Double positive (transcribing) genomes were often localized to the nuclear periphery but also appeared in the vicinity of nucleoli. To quantify the number of E1A mRNA positive genomes, we performed kymograph analysis using dual-channel t-projections of individual cells. This approach was possible because the genome position was stable over time allowing to draw a line plot through all nuclear genomes using the signal of mCherry-TAF-Iβ. When plotted as kymograph over time each genome is represented as a red line. Genomes positive for the MS2BP-NLS-GFP are represented as a green line and were scored as transcribing genomes (Fig 3B). Kymograph analysis of several cells showed that ~20% of detected genomes inside a given nucleus are positive for MS2BP-NLS-GFP (Fig 3C), which was very similar to the RNAscope approach (Fig 3D). Taken together, the analysis of transcriptionally active genomes in living cells supports our findings that early in infection transcription occurs from a subpopulation of imported genomes that probably require to actively remodel the viral chromatin for transcription onset. Longer observations of individual nuclei (10 min) confirmed the existence of transcribing and non-transcribing genomes with no indication of transcriptional alternation between genomes (Fig EV2C).

Chromatin architecture of the invading Ad genome

Transcripts accumulated with a delay of 1 h after nuclear entry, opening the question of whether structural changes may be required for assembly of the transcriptional machinery. To address the Ad pVII nucleoprotein organization, we partially digested viral and cellular genomic DNA with the endo-nuclease MNase (Fig 4A). This enzyme preferentially hydrolyses DNA in the linker region between nucleosomes and is restricted by stable protein-DNA complexes (Noll *et al*, 1975). Incubation of cellular chromatin with increasing concentration of MNase resulted in the appearance of a typical MNase ladder, reflecting the monomer and multimers of the nucleosomal DNA fragments (Fig 4A). MNase concentrations were titrated such that one sample exhibited less than 20% mono-nucleosomal DNA (low digestion) and the other sample was fully hydrolyzed to the mono-nucleosomal DNA (high digestion; Appendix Fig S2A). The latter is the common condition used for MNase-seq experiments in order to study nucleosome positions (Schones *et al*, 2008; Valouev *et al*, 2011; Diermeier *et al*, 2014), whereas approaches using limited MNase concentrations have been used to analyze chromatin accessibility (Mieczkowski *et al*, 2016; Schwartz *et al*, 2018; Hu *et al*, 2019). High and low digestion conditions were used at each infection time point to isolate the mono-nucleosomal DNA band from two replicates. The nucleosomal DNA of 17 individual experiments was used for library preparation and paired-end high-throughput sequencing, yielding in total 3.8 billion paired-end reads (Dataset EV2).

DNA fragments were aligned to the human and the Ad genome, first analyzing the read length distributions on both genomes

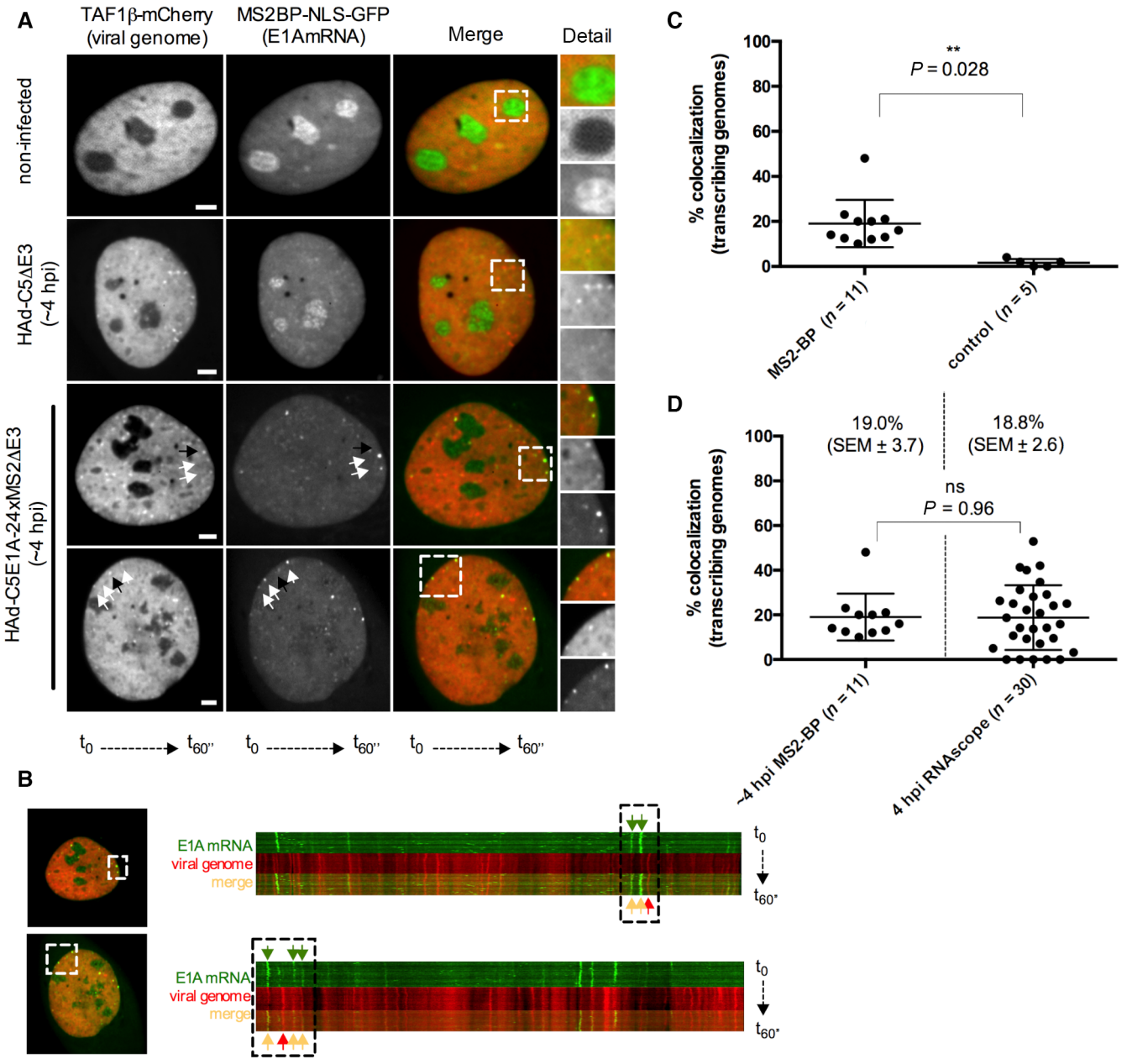

Figure 3. E1A transcription in living cells.

A Representative images showing t-projections of 60 frames (frame rate 1F/s) imaged at ~4 hpi. Cells express mCherry-TAF-Iβ showing viral genomes (first column, red signal in merge) and MS2BP-NLS-GFP showing MS2-repeat tagged E1AmRNAs (second column, green signal in merge) and an overlay (third column, merge) of both channels. The first row shows non-infected control cells, the second row cells infected with untagged control virus and the third and fourth row examples of cells infected with virus expressing E1AmRNA tagged with MS2-repeats. Black arrows point at non-transcribing and white arrows at transcribing genomes. Details are shown as magnifications of the boxed area. Scale bar is 10 μm.

B Kymograph analysis of the bottom two cells shown in (A) and as described in the material and methods section. Cell overview to the left, kymograph separated into the individual channels and a superimposition of both channels to the right. Genome position over time appears as red vertical line, E1AmRNA signal as green vertical line. The boxed area in the overview corresponds to the boxed area in the kymograph. Green arrows point at the E1AmRNA signal, orange arrows at transcribing (double positive) and red arrow at non-transcribing (single positive) genomes.

C Quantitative kymograph analysis. Scatter plot representation of the percentage of MS2-BP positive genomes per cell (n = 11 cells) versus control cells infected with viruses expressing untagged E1AmRNA (n = 5 cells) determined by kymograph analysis. Error bars represent mean ± SD.

D Scatter plot representation comparing the percentage of E1AmRNA positive genomes per cell determined with the MS2-BP method (left, n = 11 cells) versus the RNAscope method (right, n = 30 cells). Error bars represent mean ± SD.

Data information: (C, D) Unpaired t-test was used for statistical analysis. **$P \leq 0.05$; ns = not significant.
Source data are available online for this figure.

(Fig 4B). For the initial viral chromatin analysis, we chose the samples at time point 0 hpi. At this stage of infection, the genome has not been released from the capsid (Fig EV1A), and therefore, the MNase digestion pattern displays the nucleoprotein organization of the virus particles. When comparing the virus and host genomes digested at the high MNase concentration, we observed a clear difference in the read-length distribution on the two genomes (Fig 4B). Whereas most DNA fragments exhibited the typical nucleosomal DNA size on the human genome (147 bp), the protected DNA fragments on the Ad genome were significantly shorter with a mean size of 60–70 bp. The MNase-seq approach revealed a stable adenoviral packaging unit that most probably represents pVII bound to DNA, which is the only abundant genome binding protein present prior to genome release from the capsid. Previous studies showed that nuclease digestion of Ad nucleoprotein resulted in discrete populations of protected DNA fragments, suggesting a chromatin-like configuration with sizes varying between 180 and 200 bp, between 30 and 90 bp, or between 30 and 60 bp depending on the conditions used (Corden *et al*, 1976; Daniell *et al*, 1981; Johnson *et al*, 2004). Assuming that the larger DNA fragments represent partial nuclease hydrolysis products, the 60–70 bp long Ad-DNA fragments mirror the previously described pVII-DNA protein complexes. Detection of these complexes (referred to as pVII peaks) allows us for the first time to analyze the occupancy and positioning of Ad pVII. We identified defined pVII peaks over the Ad genome, showing specific pVII-DNA positions on the viral DNA (Fig 4C). The MNase profiles and peak calling were highly reproducible between biological replicates at individual time points, suggesting that specific Ad genome

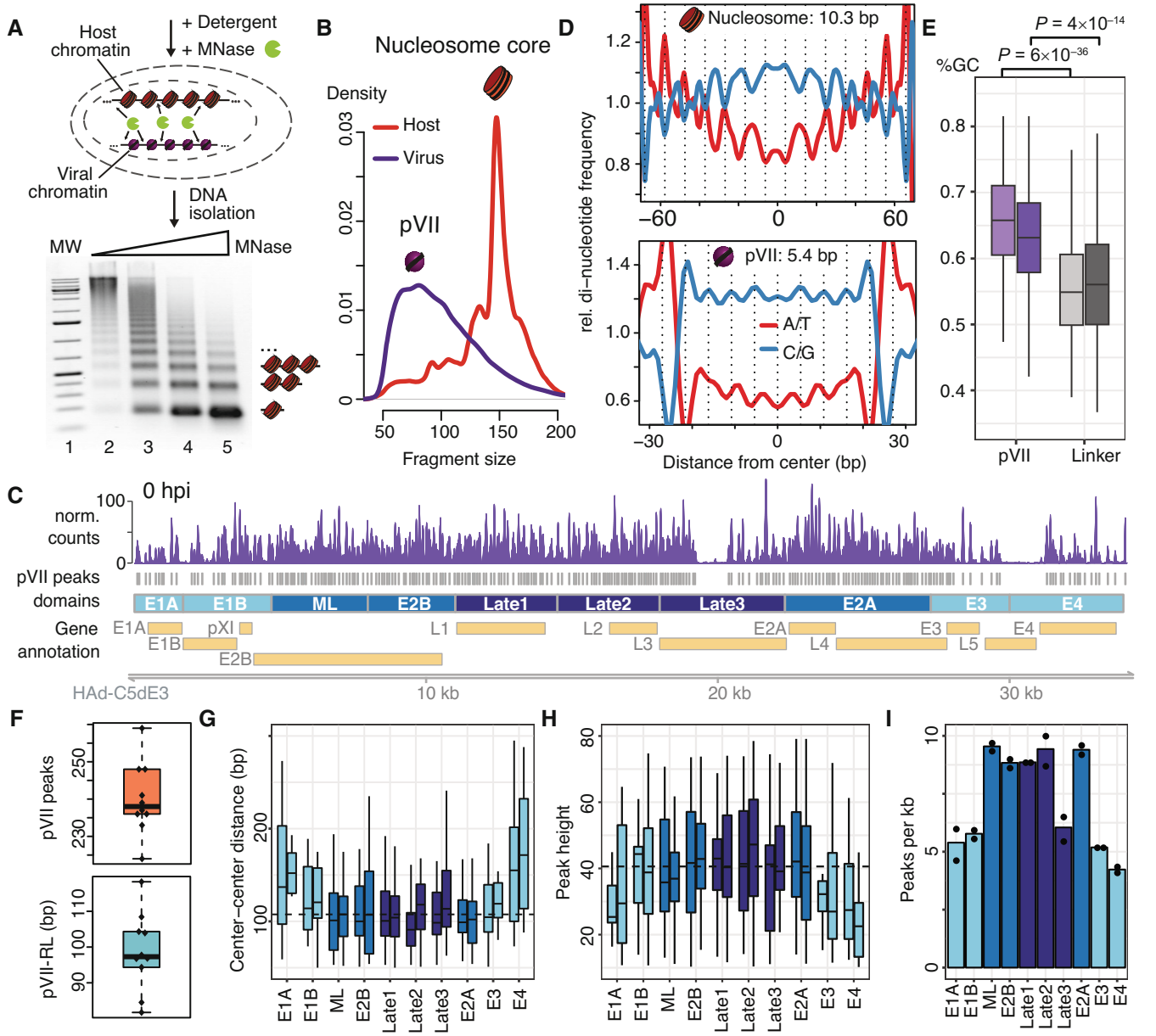

Figure 4.

Figure 4. Characterization of Ad DNA packaging by pVII.

A Schematic view of MNase digestion of infected cells. Agarose gel shows nucleosome ladders of a MNase titration. Mono-, di- and tri-nucleosomes are indicated at the right. The mono-nucleosomal DNA fraction was isolated after MNase digestions and subjected to high-throughput sequencing.

B Fragment size distribution of paired-end reads either mapped to the human genome (red) or the HAd-C5dE3 genome (purple).

C Profile illustrates HAd-C5dE3 genome coverage by MNase-seq reads. pVII positions, which have been called with the DANPOS2 package (Chen *et al*, 2013), are represented by gray boxes. Partitioning of the genome into functional domains and the timing of their transcription is indicated by the color code (lightblue = early, darkblue = late). Genomic location and underlying gene annotation of HAd-C5dE3 (NCBI accession: AY339865.1) are indicated at the bottom.

D Average WW (A/T)- and SS (C/G)-di-nucleotide frequency of nucleosomal (upper panel) and pVII fragments (lower panel). Size-selected fragments were centered at the midpoint and the di-nucleotide frequency was calculated. Every 10.3 (nucleosome)/5.4 (pVII) bp a dashed line highlights the periodicity of the signal.

E Boxplot showing GC content of pVII occupied (pVII) or free (linker) regions. Two biological replicates are shown side by side and the *P*-value of a students *t*-test of the corresponding pairs is indicated above. The GC content was assessed at 259 (replicate 1)/248 (replicate 2) pVII peaks and at 260 (replicate 1)/249 (replicate 2) linker regions.

F Boxplot showing the number of detected pVII peaks (upper panel) or the genome-wide pVII repeat length (pVII-RL) (bottom panel) for each sample (n = 10, biological replicates).

G, H Boxplots showing the distribution of pVII center-center distances (G) or peak heights (H) in each domain. Two biological replicates are shown side by side. Boxplots in (G) present n replicate 1/replicate 2 pVII center-center distances: 9/7 E1A, 17/16 E1B, 31/30 ML, 26/25 E2B, 30/30 Late1, 34/30 Late2, 27/23 Late3, 45/47 E2A, 13/13 E3, 15/16 E4. Boxplots in (H) present n replicate 1/replicate 2 pVII peak heights: 10/8 E1A, 18/17 E1B, 32/31 ML, 27/26 E2B, 31/31 Late1, 35/31 Late2, 28/24 Late3, 46/48 E2A, 14/14 E3, 16/17 E4.

I Barplot showing the number of pVII peaks per kb detected in the domains. Individual replicates (n = 2) are shown as black dots.

Data information: (E–H) Boxplots are illustrated as follows: The box represents the interquartile range (IQR). The line inside the box represents the median of the dataset. The whiskers extend from the edges of the box to the first and third quartiles, but no further than 1.5 times of the IQR.
Source data are available online for this figure.

positions are occupied with pVII, equivalent to nucleosomes on the host genome (Appendix Fig S2C–E). The Ad exhibits a defined nucleoprotein architecture that differs along its functional regions. Interestingly, a central region of the viral genome (Late3) and a region between the E3 and E4 genes exhibited almost no peaks (Fig 4C).

Regions lacking pVII peaks could result either from the absence of pVII protecting the DNA from continuous MNase degradation or from tight compaction excluding DNA from MNase digestion and therefore resulting in DNA fragments of higher molecular weight. To test both possibilities we digested Ad chromatin directly within virus particles or protein-free viral DNA (gDNA, to control for MNase cleavage preferences) with a very low amount (5 U) of MNase (Fig EV3A and B). In this assay, the isolated fragments are longer containing a mixture of mono-, di-, and multimers of pVII protected DNA, as well as un-protected DNA fragments (Fig EV3B). The level of compaction can be assessed by the local cleavage rate of MNase, which exhibits a higher cleavage rate at unprotected sites and a lower cleavage rate at sites with tight compaction. While cleavage patterns were uniform in Ad gDNA, the highest cleavage frequencies in virus particles corresponded to the regions lacking pVII signals after infection (Late3 and between E3 and E4, Fig EV3C and D), suggesting the presence of "unprotected" domains in the Ad genome, which are completely hydrolyzed under the MNase digestion conditions used in infected cells.

pVII exhibits DNA sequence dependent positioning patterns

We next asked whether intrinsic DNA sequence patterns may establish pVII peaks. Such a sequence dependent positioning code was shown for nucleosomes (Kaplan *et al*, 2009). Sequence specific bending of DNA reduces the energy requirements for wrapping the DNA around the histone octamer surface, revealed by characteristic di-nucleotide repeats in the nucleosomal DNA sequence at 10 bp intervals (Widom, 2001). This pattern was clearly observed in the human host genome (Fig 4D – upper panel).

Interestingly, we also observed a specific di-nucleotide pattern when computing the di-nucleotide frequency of the mid-point centered pVII peak fragments, suggesting that a specific DNA code underlies pVII positioning on the Ad genome (Fig 4D – lower panel). However, to our surprise, a periodic pattern of 5.4 bp was obtained for WW- (where W is A or T) and SS-dinucleotides (where S is G or C), much shorter than the nucleosomal DNA repeat pattern, implying a distinct packaging mode. Such a repeat pattern indicates a spring-like winding of the DNA around pVII. Additionally, pVII peaks are preferentially positioned at sites with higher GC content (Fig 4E). Even without knowledge of the exact DNA-pVII topology, the sequence pattern and the specific binding sites of pVII on DNA clearly suggest the organization of the Ad genome into a specific nucleoprotein-structure that may be essential for its function.

On average, we detected 238 pVII peaks covering the Ad genome, which was consistent throughout all timepoints (Fig 4F – upper panel). Next, we quantified the mean pVII-DNA particle repeat length (pVII-RL), the distances between individual pVII peaks, by measuring the center-center distance of adjacent pVII peaks in each sample (Fig 4F – lower panel and Appendix Fig S2B). The median pVII-RL measures 97.3 bp, suggesting that we isolated defined pVII-DNA particles spaced by accessible DNA linkers of 20–40 bp in length.

To address the presence of functionally distinct regions, we systematically subdivided the Ad genome into 10 domains of similar size, named after the included gene and classified by the timing of expression (color code: early to late = light to dark blue) (Fig 4C). Plotting the pVII-RL showed a clear correlation of the repeat length with expression timing (Fig 4G). Increased repeat lengths in the regions of the early genes may serve to increase genome accessibility and to enable the binding of regulatory factors. Strengthening this finding, the pVII occupancy and stability, reflected by the peak height, showed an anti-correlation with early timing in gene expression (Fig 4H). In general, less pVII peaks (Fig 4I) with larger repeat lengths and increased instability were found at early expressed domains (Appendix Fig S2D and E) indicating an open chromatin

conformation, which could be confirmed with limited MNase digestions of virus particles (Fig EV3C and D).

Our high-resolution analysis of Ad genome packaging provides, for the first time, information on pVII-DNA peak positioning and occupancy. The analysis revealed a defined viral nucleoprotein architecture, with an open configuration at early expressed genes potentially enabling immediate access of the host cell transcription machinery.

Dynamic changes of viral genome accessibility over time

One parameter determining viral genome accessibility and activity is the specific occupation with pVII, another parameter is the higher order structure of the Ad-chromatin and its global accessibility. Assessing the nucleoprotein structure with high MNase concentrations is providing information about stable protein particles occupying the DNA (Schwartz *et al*, 2018). In contrast, limited MNase concentrations, in analogy to the analysis of regulatory sites by DNase I hypersensitivity mapping (Längst *et al*, 1997), can reveal accessible genomic regions and regulatory elements (Schwartz *et al*, 2018).

Only a fraction of about 20% of the viral genomes in the nucleus are transcriptionally active (Figs 2E and 3C), explaining why we observed only little variation between the different time points treated with high MNase concentrations (Fig 5A). In contrast, at low MNase concentrations (Fig 5B and C) clear changes over infection time are detected, signifying gross structural changes of the viral chromatin during early infection. We hypothesize, that the active viruses undergo de-compaction that can be detected by limited MNase hydrolysis. When monitoring the pVII peak positioning at 0 hpi, the low MNase signal was well correlated with the high MNase signal ($R = 0.91$), essentially showing the same pVII peak positions in both conditions (Fig EV4A and B). However, when inspecting the viral chromatin maps at later time points, we detected specific structural changes, with changes in peak intensities and the appearance of novel peaks in the low MNase signal (Fig 5C). To identify structural changes over time, we quantified the release rates of pVII from Ad chromatin in low MNase digestions. A similar strategy was used in a previous study to measure chromatin accessibility dynamics (Mieczkowski *et al*, 2016; Mueller *et al*, 2017; Chereji *et al*, 2019). The average signal intensity at each pVII position was quantified and the log(fold-change) against 0 hpi was used as a score for chromatin accessibility changes (pVII accessibility score; Fig 5C). As negative accessibility scores are a result of the normalization of the signal and not due to further compaction of Ad chromatin compared to 0 hpi, we considered only positive scores representing opening of genomic regions and being more MNase accessible.

The increase of Ad chromatin accessibility is correlating with viral transcriptional activity, with Ad chromatin remodeling and de-condensation at the early gene regions (Fig 5D). Interestingly, chromatin opening at E1A is rapid and already decreasing at 4 hpi. As this is the first gene to be transcribed, we suggest that this region already exists in an open chromatin configuration, correlating with the pre-existing low pVII peak density, and therefore additional opening is not that pronounced as at other early genes over time.

The Late3 domain also exhibited a positive accessibility score, suggesting Ad chromatin de-condensation at this site (Fig 5C). Additionally, in this domain, we observe a discrete peak appearing at low MNase concentrations at 4 hpi (Fig 5C shaded region and peak marked by asterik; Fig EV4C). The occurrence of this local nuclease hypersensitive site suggests the binding of regulatory factors (Kent *et al*, 2011; Mieczkowski *et al*, 2016; Chereji *et al*, 2017). Transcription factor motif analysis revealed potential binding sites for EN1, CREM, C/EBP-alpha, and C/EBP-beta in the center of the peak (Fig EV4C). Interestingly, the transcripts of CREM, and C/EBP-beta in the host cell were significantly upregulated up to twofold upon viral infection (Fig EV4D).

In summary, the time-resolved chromatin accessibility analysis shows, that in addition to the predefined pVII binding sites on the genome, Ad chromatin compaction specifically changes over time. Chromatin opening of the early gene regions starts at 0.5 hpi, prior to transcriptional activation. We suggest that changes in early gene regions are a pre-requisite for gene activation process.

Dynamic changes of Ad genome packaging

Previous studies showed that histones are deposited onto the Ad genome (Komatsu *et al*, 2011; Komatsu & Nagata, 2012; Giberson *et al*, 2018). However, at which position and whether these are functional nucleosomes is not known.

Analyzing the MNase-seq data of the time points 1, 2, and 4 hpi revealed the appearance of 147 bp long, protected DNA fragments at 2 and 4 hpi, after nuclear import of the viruses and prior to early gene activation (Fig 6A). The nucleosome sized peaks increased in intensity, within the background of the more abundant pVII protected DNA fragments, suggesting co-occupation of the viral genome by pVII and nucleosomes. To quantify the fraction of nucleosomes on the viral genome, we simulated the peak sizes by computationally adding nucleosomal reads to the pVII fragment data of time point 1 hpi (Fig EV5A). The analysis suggested that at 2 hpi 8% and at 4 hpi 14% of the viral genomes are occupied with nucleosomes.

To verify this indirect prediction of nucleosomal occupancy on the Ad genome, we generated U2OS cell lines, stably expressing FLAG-tagged histone H3.1 (U2OS31) and the histone variant H3.3 (U2OS33) (Fig EV5B and C). Stable cell lines were infected with Ad and ChIP-seq experiments were performed at 0, 1, 2, and 4 hpi, using either FLAG- or H3K27ac-antibody.

At 0 hpi we barely detected any ChIP-seq reads mapping to the Ad genome, showing the high specificity of our ChIP-seq approach (Fig 6B). The number of precipitated Ad DNA fragments increased with infection time for histone H3.3 (FLAG IP: U2OS33), in contrast to the canonical histone H3.1 DNA fragments (FLAG IP: U2O31) (Fig 6B). The low number of histone H3.1 DNA fragments clustered with the samples taken at 0 hpi on PCA and hierarchical clustering analysis (Figs 6C and EV5D), and did not show any specific sites of enrichment on the Ad genome (Fig EV5E, upper panel), suggesting no specific recruitment of histone H3.1. However, we detected specific H3.3 binding sites on the Ad genome, clearly showing that this histone variant is exclusively deposited onto incoming viral genomes (Fig 6D). The H3.3 ChIP-Signal was most pronounced at early transcribed gene promoters, including E1A, E3, and E4, but not E1B. Nevertheless, H3.3 also accumulated at additional loci throughout the Ad genome, like the E2B and Late2 regions (Fig 6D and E).

Next, we performed H3K27ac ChIP-seq experiments, showing efficient histone acetylation, without any detectable delay after

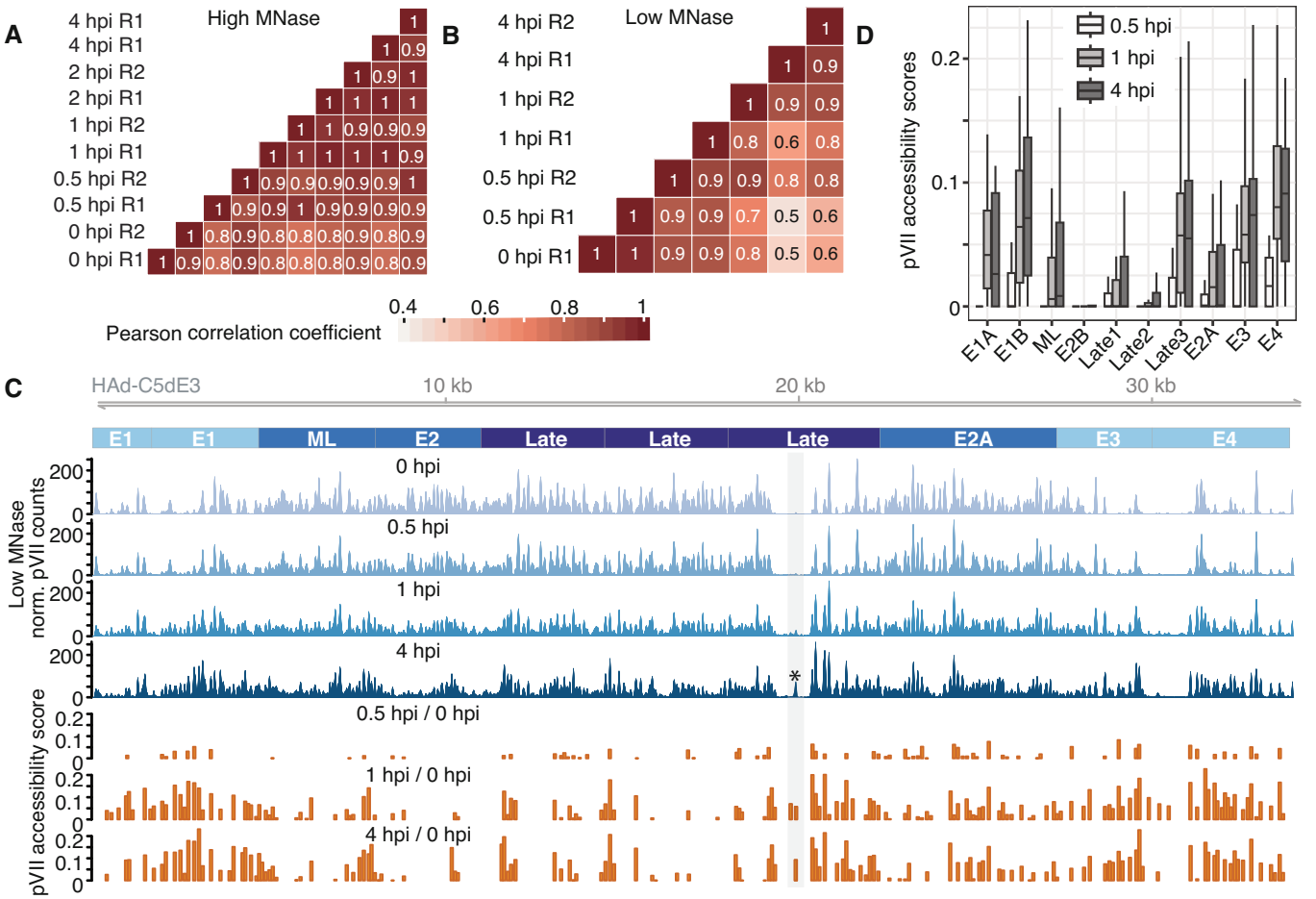

Figure 5. Early decondensation of Ad chromatin.

A, B Sample-wise Pearson correlations of (A) high MNase or (B) low MNase profiles.

C Profile illustrates HAd-C5dE3 genome coverage by low MNase-seq fragments. The average of two replicates is shown, except at timepoint 0 hpi where only one replicate was available. The accessibility score was calculated as the log(fold-change) between the indicated timepoint and 0 hpi. The score was assessed for each pVII peak (orange bars) and negative scores were set to 0. A new accessibility peak arising during infection in the Late3 region is marked by an asterisk.

D Boxplot showing the accessibility score distribution in each domain at each tested timepoint after infection. The box represents the interquartile range (IQR). The line inside the box represents the median of the dataset. The whiskers extend from the edges of the box to the first and third quartiles, but no further than 1.5 times of the IQR. Boxplots represent n pVII accessibility scores at 0.5/1/4 hpi in the respective domain: 6/11/7 E1A, 14/22/23 E1B, 38/43/41 E2A, 23/23/23 E2B, 12/17/17 E3, 13/22/20 E4, 25/28/29 Late1, 26/27/27 Late2, 23/29/31 Late3, 24/25/26 ML.

Source data are available online for this figure.

histone deposition on the Ad genome. Interestingly, we observed histone acetylation at specific sites on viral genomes, representing a subset of all H3.3 binding sites, marking these as active regions (Fig 6B and D). As the H3K27ac antibody does not discriminate between H3.1 and H3.3, the results obtained in U2OS31 and U2OS33 are very similar (Figs 6B and C, and EV5D). Histone acetylation was identified at all early gene promoters, except E1B, with the strongest signal at E1A, which was already detectable 1 hpi (Fig 6D and F). Additionally, we observed pronounced K27 acetylation at two central regions covering the E2B or Late2 gene.

We expected that sites marked by H3K27ac would promote chromatin de-condensation. However, neither H3.3 deposition nor H3K27 acetylation did correlate with chromatin accessibility changes ($R^2 < 0.1$, Appendix Fig S3A and B). A closer inspection showed that acetylation at the E1A and E4 genes was associated with open chromatin (Appendix Fig S3C and F), whereas the acetylated Late2 and E2B sites in the genome center remain condensed (Appendix Fig S3D and E).

In summary, we suggest that the histone variant H3.3 is preferentially deposited onto viral genomes and a subset of binding sites get acetylated at K27. H3.3 acetylation is detected preferentially at early gene promoters, being nucleosomal and acetylated prior to transcription initiation.

Early genes gain a positioned +1 nucleosome

The sonicated DNA in the ChIP experiments is not short enough to reveal the exact sites of nucleosome assembly. Therefore, we

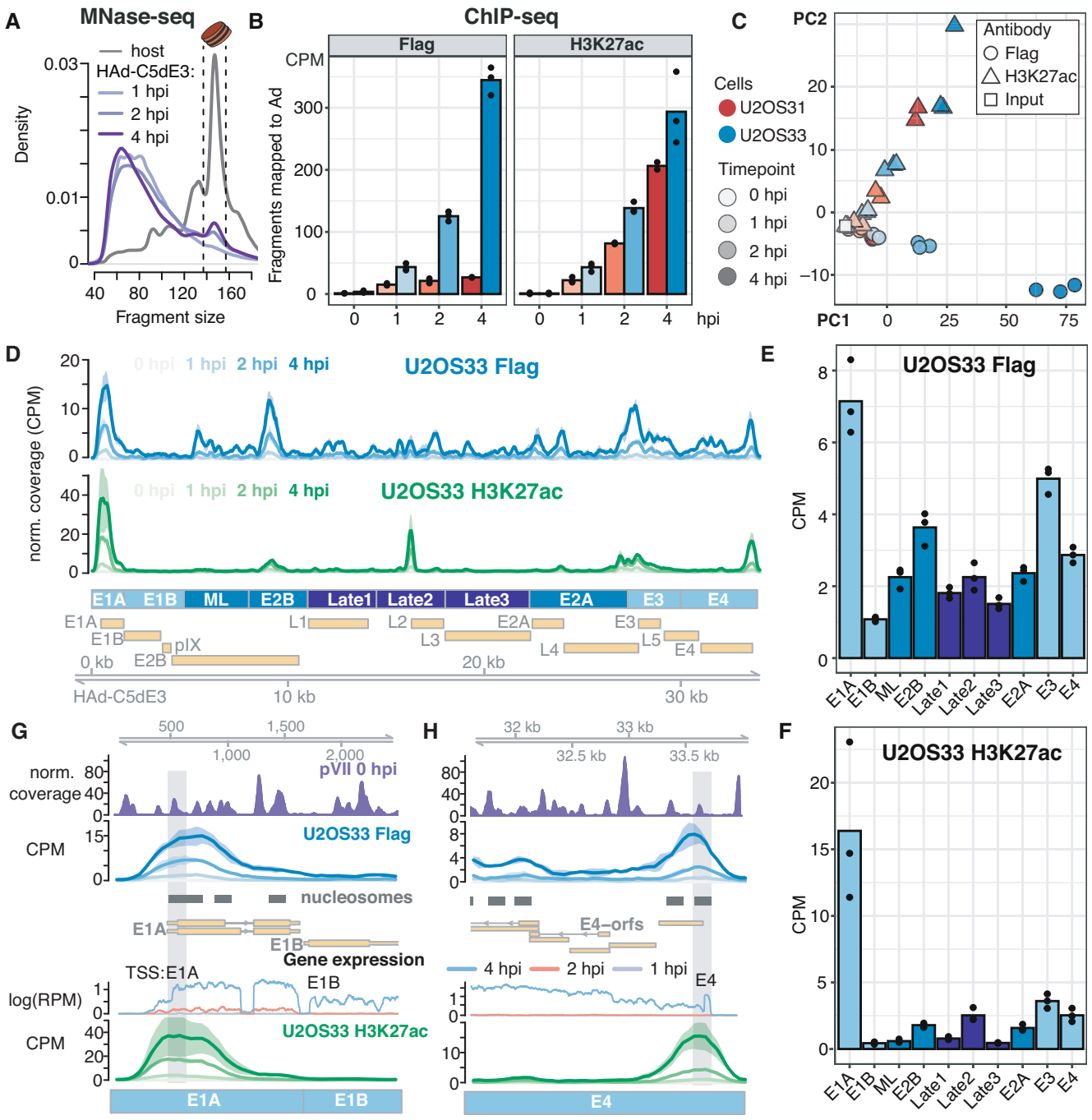

Figure 6. Nucleosome assembly onto the HAd-C5dE3 genome.

A Fragment size distribution of paired-end reads from high MNase experiments (see Fig 4) either mapping to the human genome (gray) or the HAd-C5dE3 genome (shades of purple). Nucleosome sized fragments are highlighted between two dashed lines at 137 and 157 bp.

B Quantification of ChIP-seq fragments mapping to the HAd-C5dE3 genome over time. Flag/H3K27ac immunoprecipitation (IP) was conducted in U2OS cells expressing either Flag-tagged H3.1 (U2OS31, blue shades) or H3.3 (U2OS33, red shades). The barplots show the average of 2–3 replicates (black dots) of each time point and condition. The antibody used in the IP is indicated at the top.

C Principal component analysis (PCA) of ChIP-seq samples.

D H3.3 and H3K27ac profiles along the Ad genome. Normalized coverage of Flag (blue shades)/H3K27ac (green shades) ChIP-seq experiments in U2OS33. The average signal of 3 replicates is shown and the 95% confidence interval is indicated. CPM, counts per million mapped fragments.

E, F Average coverage of Flag-tagged H3.3 (E) or H3K27ac (F) in specific Ad domains.

G, H Genome browser tracks showing pVII occupancy at 0 hpi (upper panel), the H3.3 (blue shades)/H3K27ac (green shades) accumulation over time, predicted sites of nucleosome assembly (gray boxes) using nucleosome-sized fragments (137–157 bp) from 4 hpi MNase-seq samples and expression levels of early genes (middle panel).

combined the ChIP-seq data with the MNase-seq derived nucleosomal DNA fragments (size between 137 and 157 bp; time point 4 hpi), allowing us to reveal the exact positions of the nucleosome core particle. A cut-off of 2 CPM was used to define the genomic sites of preferential nucleosome assembly from U2OS2 FLAG IPs (Appendix Fig S3G, blue boxes) and the exact nucleosome positions were determined using MNase size selected fragments (Appendix Fig S3G, golden boxes). The nucleosome profile was distinct from the pVII landscape and nucleosome calling was reproducible between replicates.

Nucleosome assembly occurs in the genomic regions associated with the early genes, exhibiting a lower pVII peak occupancy and increased pVII peak repeat-length. An analysis of the incoming nucleosome positions with respect to pVII peak occupancy revealed that nucleosomes directly formed over the central positions of the pVII-nucleoprotein particles (Appendix Fig S3H) suggesting a specific mechanism to replace pVII.

A detailed analysis of the nucleosome assembly sites with respect to the early genes revealed that nucleosomes occupy discrete positions. Remarkably nucleosome positions coincide with the 5′-end of the synthesized E1A, E4, E3, and E2A RNA, correlating with the highly positioned +1 nucleosome found in active human genes (Fig 6G and H and Appendix Fig S3I) (Schones *et al*, 2008). Additional nucleosomes are detected in the coding region of the early genes. The +1 nucleosome marks the transcription start sites of genes and flanks the regulatory sequences of the promoters, residing in the upstream DNA linker region.

The time-resolved chromatin analysis shows, that the predefined pVII chromatin organization is remodeled, exchanging pVII for histone H3.3. Histone H3.3 deposition and acetylation occur mainly at the +1 region of early genes and happens prior to transcription onset. Our results imply the requirement of specific nucleosome assembly for transcription initiation at early genes.

Discussion

The adenovirus exhibits a defined nucleoprotein architecture

Early studies suggested that pVII binds adenoviral DNA at a 1:1 ratio like histones and protects DNA from endonuclease cleavage (Lischwe & Sung, 1977; Sato & Hosokawa, 1984). However, no defined pVII-DNA particle has been identified on functional Ad genomes yet, as the protected DNA had a variable size range of 30–250 bp depending on experimental conditions and the degree of endonuclease digestion (Corden *et al*, 1976; Daniell *et al*, 1981; Mirza & Weber, 1982). In addition, it is still not known whether pVII binds to specific sites on the viral genome, creating a defined viral nucleoprotein architecture, like the cellular nucleosomes. Here, our MNase-seq approach digesting Ad together with the host chromatin at different time points post infection, revealed the 147 bp long DNA fragments of the host nucleosomes and identified a shorter, protected DNA species with a mean size of about 60–70 bp originating from the viral sequences. With pVII being the major DNA binding protein of the adenovirus, and the protected fragment species being present prior to genome release from the capsid, the observed DNA fragments are very likely protected by their association with pVII. The viral MNase digestion pattern is not as regular as the nucleosomal DNA fragments, suggesting heterogeneity in pVII-DNA complexes, as observed in the early studies. Such a digestion pattern could also be the result of a discrete pVII core complex (60 bp) that tightly associates with neighboring pVII complexes, partially protecting the interconnecting linker DNA from endonuclease cleavage. Such structures were indeed visible in early electron microscopic studies (Mirza & Weber, 1982).

pVII organizes the viral DNA into a specific nucleoprotein structure

Annotation of the putative pVII protected DNA fragments to the viral genome revealed for the first time a specific positioning of pVII peaks on the genome. We show that the Ad genome is covered by an average of 238 defined pVII peaks located at defined positions that are spaced by DNA linkers of about 20–40 bp. Precise pVII peak positions suggest a highly organized and defined nucleoprotein architecture shared by all virus particles. In host cells, DNA structure determines histone binding preferences, as shown by the DNA sequence dependent nucleosome positioning code (Kaplan *et al*, 2009). DNA bending allows wrapping of the DNA around a protein core, requiring site specific SS and WW dinucleotide enrichments. SS and WW repeats differ in the wideness of the DNA major/minor groove thereby bending the DNA (Kaplan *et al*, 2009). If the sequence dependent DNA shape follows the protein surface path it increases binding affinity and thereby determines binding site preferences (Widom, 2001). Whereas the nucleosomal di-nucleotide repeats are spaced by 10.3 bp, pVII exhibits a much shorter repeat length of 5.4 bp. Nucleosomes and pVII-DNA complexes are organized differently, and we suggest that the DNA associates in form of a spiral around the pVII core (Fig 4D).

Our results suggest that pVII folds Ad genomes into an organized higher order structure, driven by specific binding sites predetermined by the Ad DNA sequence. A recent report showed that pVII is not required for genome packaging into virions, albeit the lack of pVII resulted in non-infectious Ad (Ostapchuk *et al*, 2017). The specific organization of the viral genome with pVII may be functionally linked with endosomal escape, in addition to potential gene regulatory effects. However, as endosomal escape of the Ad lacking pVII is blocked and genomes are not delivered to the nucleus, the effect of pVII absence on gene expression and the specific exchange of positioned pVII peaks by nucleosomes cannot be concluded. Our observation that pVII peaks are reproducible across biological replicates strongly indicates a specific positioning code. Protein VII peak densities and the changes in pVII occupancy upon nuclear delivery (see below) correlate with the onset of the transcriptional program and cellular factors, such as TAF-Iβ. TAF-Iβ was shown to positively influence viral transcription, being recruited to the incoming genomes via pVII (Matsumoto *et al*, 1995; Haruki *et al*, 2003). One possible approach to validate our findings of an underlying sequence code for pVII positioning may be the rational (sequence-)design of therapeutic DNAs packed into adenoviral vectors to achieve high and sustained transgene expression. *In vitro* assembly of pVII-DNA complexes showed that pVII binds to DNA regardless of the DNA sequence (Komatsu *et al*, 2011; Avgousti *et al*, 2016). However, it was not addressed, whether sequence motifs are preferred by pVII, like it was observed in

nucleosome assemblies. Specific assembly factors and ATP dependent remodeling enzymes may be required to adjust the precise positioning of pVII on DNA, in addition to the DNA assembly code. Like nucleosomes, pVII does not bind as isolated protein to DNA but is forming homotypic complexes. The number of pVII proteins per Ad genome was estimated to be about 500 copies (Benevento *et al*, 2014) and early biochemical studies suggested the occurrence of pVII dimers (Everitt *et al*, 1975; Chatterjee *et al*, 1985). Taking into account the median number of 238 pVII positions we identified on the Ad genome (plus 19 on average for the 1,871 bp deleted E3 region) our data would be compatible with two pVII molecules forming dimers at 238 discrete sites. This number would also come close to the ~200 "beads-on-string" estimated for purified viral cores (Vayda *et al*, 1983).

Viral chromatin dynamics drives early gene expression

Here we show that nuclear delivery of viral genomes takes about 30 min but requires an additional hour before viral transcripts appear. In agreement with earlier observations, full activation of immediate early gene expression is reached after 4 h (reviewed in Pied & Wodrich, 2019). Correlating gene expression with our differential MNase digestion conditions, we show that early genes exhibit a reduced density of MNase protected regions that most probably represent pVII peaks, and increased pVII peak repeat lengths. Direct detection of genomes and of the E1A mRNA showed that genomes imported to the nucleus require about 2 h to activate gene expression and to synthesize the mRNA (i.e. E1A transcription). In this lag phase, pVII is selectively replaced by nucleosomes at early genes, preceding gene activation. ChIP-seq experiments revealed that the histone variant H3.3 is incorporated at many specific genomic locations on the Ad genome, including the early gene regions, shortly after nuclear import. A subset of these nucleosomal sites, corresponding mainly to the early gene positions, are acetylated at H3K27 prior to transcriptional activation. We identified that 80% of the imported Ad genomes are transcriptionally inactive, but the remaining 20% do actively rearrange the chromatin structure to allow transcription factor and RNA Polymerase binding. Once entering the nucleus, we showed that the chromatin structure of active viral genomes decondensed and viral DNA becomes more accessible. A delay for E1A gene expression and heterogeneity of transcriptional activity for imported genomes was recently reported supporting our findings (Suomalainen *et al*, 2020). The limited activation of incoming genomes and subsequent de-compaction was also observed during herpes simplex virus type 1 (HSV-1) infections showing common features of nuclear replicating DNA viruses (Sekine *et al*, 2017; Hu *et al*, 2019). However, in the case of HSV-1 gene activation, it was suggested that the accessibility over the complete viral genome increased (Hu *et al*, 2019), whereas active Ad genomes selectively de-condense the early gene regions and a specific domain in the Late3 gene. The overlap of open viral chromatin with histone H3K27 acetylation and early transcription, suggests an organization that allows the binding of regulatory factors for gene activation. In contrast, for the Late3 region no functionality in gene expression control has been assigned yet. Our observation that increased DNA accessibility overlaps with binding motifs for transcriptional regulators may indicate a role as a transcriptional enhancer.

pVII is exchanged for nucleosomes at early genes

Genome de-compaction correlates with low pVII occupancy and coincided with a site-specific exchange of pVII protein for histone octamers. The characteristic nucleosomal fragment length of 147 bp, observed in MNase-seq, confirms the release of pVII from the viral genome, as pVII-nucleosome complexes would result in longer fragments after MNase digestion as previously shown (Avgousti *et al*, 2016). The ChIP-seq experiments clearly show an overlap of the 147 bp long protected DNA fragments and the location of the histone variant H3.3. Furthermore, the newly assembled nucleosomes are located on previous positions of pVII-DNA complexes. Once pVII is set free, it is able to sequester HMGB1 and modulate the cellular chromatin structure, thereby downregulating the immune response of the host cell (Avgousti *et al*, 2016). Protein VII molecules were exchanged for histone octamers downstream of the transcription start site, overlapping the +1 nucleosome site, of actively transcribed genes (i.e. the E1A and E4 gene, Fig 6G and H). The calculated and experimentally validated number of nucleosomes and their positions on DNA, as determined by the 147 bp long DNA fragments and the ChIP-seq experiments, corresponds to the number of ~20% transcribing genomes, exhibiting single nucleosomes positioned at the +1 position of early genes. As nucleosome exchange is commencing at 0.5 hpi and is completed at 2 hpi when transcription activation is detectable by mature transcript accumulation, we suggest that nucleosome assembly and positioning is replication- and transcription-independent, being a prerequisite for gene activation. At this point, we can only speculate why only a subpopulation of genomes is transcriptionally active. The reasons could be that the few genomes undergoing nucleosome assembly and active transcription produce the replication enzymes, whereas the bulk of genomes enters replication without activation as an elegant way to avoid repeated chromatinization. Alternatively, cellular defense mechanisms and/or the cell cycle could influence transcription levels (Suomalainen *et al*, 2020). The heterogeneity of 80% inactive genomes and 20% activated genomes complicates the analysis of the MNase-seq data. High MNase concentrations do not differentiate between both states, and we suggest that low MNase conditions capture the dynamic viral proportion, changing and preparing its genome for gene activation. The data nicely suggest such a scenario, but there is the caveat that we catch an effect of the mixed population that we cannot differentiate. Out of necessity to reach the desired sequencing depth and to be consistent in our comparative experimental approach we mainly explored infections using relative high MOI. However, we did observe transcriptionally active genomes also at lower MOI (Fig EV2B). In addition, many but not all transcriptionally active genomes are located to the nuclear periphery, suggesting that dose and positional effects may also influence transcriptional activation. This could be especially true for very low MOI during natural infections. The experimental toolbox developed in this study will be an asset to address such pertinent questions in the future.

Interestingly, pVII removal and nucleosome placement at early genes preceded transcription, suggesting that pVII removal is driving transcription and not vice versa, as previously reported (Chen *et al*, 2007). This observation is in agreement with recent pVII-ChIP experiments showing transcription and replication independent pVII removal and nucleosome depositioning in early infection (Komatsu *et al*, 2011; Komatsu & Nagata, 2012; Giberson *et al*, 2018).

Future work has to reveal which machinery drives the pVII nucleosome exchange and determines specificity. Generally, active promoters are free of nucleosomes (nucleosome depleted region, NDR) and the transcription start site is typically located upstream of a well-positioned "+1" nucleosome influencing recruitment of the transcriptional machinery and initiation of transcription (Jiang & Pugh, 2009; Kubik *et al*, 2015). In the case of the Ad genome, our data show that the E1A promoter region is already pVII free and exists in an accessible nucleoprotein conformation when entering the cell, potentially allowing fast assembly of the +1 nucleosome and enabling immediate gene activation. This organization might thus be defined in the producer cell during genome packaging, which occurs in a sequential packaging sequence driven manner (Hammarskjöld & Winberg, 1980; Hearing *et al*, 1987). In contrast, other early regions were actively de-compacted with nuclear entry of the virus, correlating with slightly delayed transcriptional onset. This effect may be driven by pVII interaction with transcription promoting factors like TAF-Iβ and E1A (Johnson *et al*, 2004; Haruki *et al*, 2006). Besides pVII is post-translationally poly(ADP)-ribosylated or acetylated in infected cells, which might further promote chromatin relaxation (Déry *et al*, 1986; Avgousti *et al*, 2016).

Taken together, our analysis provides unprecedented insight into the adenoviral nucleoprotein structure and its dynamic reorganization during the onset of viral gene expression. We reveal genome sequence, pVII-DNA occupancy and nucleosome positioning as driving forces for transcriptional activation. The structural analysis is still limited to a single adenovirus genotype and it will be interesting to test whether these dynamic changes are conserved among other adenoviruses. Furthermore, reproducing such organization in adenoviral vectors could result in efficient and sustained transgene expression.

Materials and Methods

RNA library preparation and sequencing

Total cellular RNA was isolated from infected H1299 cells at 0, 0.5, 1, 2, 4 hpi using TRI Reagent (Molecular Research Center, Inc.). Two time series sets of total RNA were prepared from two independent infections. Library preparation, comprising rRNA depletion, polyA enrichment, fragmentation to ~270 nucleotide length, reverse transcription to cDNA using random hexamer primers and adapter ligation, was accomplished using TruSeq RNA Sample Preparation Kit v2 (Illumina) according to the manufacturers protocol by EMBL GeneCore facility in Heidelberg. Libraries were sequenced on Illumina HiSeq2000 platform resulting in 18–75 Mio 50 bp single-end reads per sample.

Differential MNase-seq, library preparation, and sequencing

H1299 cells infected with Ad were permeabilized with the detergent IGEPAL-CA630 and digested at 0, 0.5, 1, 2 and 4 hpi with 100 U/600 U of MNase for 4/5 min at 37°, denoted as "low"/"high" MNase. Mnase treatment was performed as described in (Schwartz *et al*, 2018). After fragmented DNA has been separated on 1.3% agarose gels, mono-nucleosomal fractions were isolated. Paired-end sequencing on Illumina HiSeq2000 platform (EMBL GeneCore facility, Heidelberg), yielding about 220 million read pairs of 50 bp length per sample, was accomplished following library preparation using NEBNext DNA library prep Master Mix Kit (New England Bioloabs).

ChIP-seq, library preparation, and sequencing

Chromatin immunoprecipitation (ChIP) was performed as described previously with slight modifications (Minderjahn *et al*, 2020). Chromatin of cell lines encoding Flag-tagged H3.1 or H3.3 (U2OS31 and U2OS33, respectively) was harvested at 0, 0.5, 1, 2 and 4 hpi in triplicates of independent Ad infections. Briefly, for anti-H3K27ac and anti-Flag ChIP-seq, cells were crosslinked with 1% formaldehyde for 10 min at room temperature and the reaction was quenched with glycine at a final concentration of 0.125 M. Chromatin was sheared using ultrasonication (Covaris S2). A total of 2.5 µg of antibody against H3K27ac (Abcam, ab4729), or FLAG (M2, Sigma Aldrich, F3165), was bound to 20 µl pre-washed Dynabeads Protein A (Thermo Fisher Scientific) in a total volume of 200 µl PBS containing 0.02% Tween-20 for 1 h at room temperature rotating at 6 rpm. The buffer was removed on a magnet and antibody-coupled beads were resuspended in 120 µl dilution buffer (20 mM Tris pH7.4, 2 mM EDTA, 100 mM NaCl, 0.5% Triton x-100) containing proteinase inhibitors (cOmplete™, EDTA-free Protease Inhibitor Cocktail, 50×, Sigma). 80 µl sonicated chromatin was added to the antibody-coupled Dynabeads (sonicated chromatin of approx. 1.7×10^6 cells) and incubated for 3 h at room temperature rotating at 6 rpm. Beads were washed on a magnet and chromatin was eluted. After crosslink reversal, RNase A and proteinase K treatment, DNA was extracted with the Monarch PCR & DNA Cleanup kit (NEB). Sequencing libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB) according to manufacturer's instructions. The quality of dsDNA libraries was analyzed using the High Sensitivity D1000 ScreenTape Kit (Agilent) and concentrations were assessed with the Qubit dsDNA HS Kit (Thermo Fisher Scientific). Libraries were sequenced paired end on a NextSeq550 (Illumina).

Viruses and virus production

Bacterial artificial chromosome (BAC) mutagenesis was used to generate recombinant adenoviruses following the protocol described in (Ruzsics *et al*, 2014) and (Martinez *et al*, 2015). Briefly, bp 1–3,513 of the HAdV-C5 genome including the left inverted terminal repeat (ITR) and the whole E1 region was replaced by an FRT (Flippase Recognition Target) site and maintained as BAC (termed B12) to perform lambda red recombination essentially as described. Plasmid pO6-Ad5-E1-FRT (encoding the left ITR and the complete E1 region followed by an FRT recombination site) was used to clone 24 copies of the MS2 binding site repeats into the 3′UTR of the E1A gene. Subsequently, flp-recombination was used (Ruzsics *et al*, 2014) to recombine plasmids pO6-Ad5-E1-FRT-MS2 and BAC B12 to reconstitute BxAd5-E1-MS2 a BAC encoding the whole AdV genome including the E1A gene tagged with 24 copies of the MS2 binding site except a short deletion in the E3 region. To generate the virus the BxAd5-E1-MS2 BAC was digested by *PacI* restriction enzyme to release the viral genome and transfected into HEK293 cells to reconstitute the recombinant viruses (termed in this manuscript

HAd-C5E1A-24xMS2dE3). Two rounds of plaque purification were performed followed by large scale production in HEK293 cells and purification through double $CsCl_2$ gradients banding. Replicative E3-deleted HAd-C5dE3 virus or E1/E3-deleted HAd-C5dE3 GFP expressing control virus were essentially produced in the same way. Purified viruses were extensively dialyzed against PBS/10% glycerol and stored in aliquots at −80°C and quantified as physical particles per µl following the method described in (Mittereder *et al*, 1996).

Cell culture and infection

U2OS (ATCC #HTB-96) and H1299 (ATCC #CRL-5803) and HEK293 (kindly provided by G. Nemerow, TSRI, La Jolla, USA) cells were maintained at 37°C and 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM)-Glutamax supplemented with 10% fetal calf serum (FCS) and 1% of Penicillin/Streptavidin. U2OS cells stably expressing mCherry-TAF-Iβ (Komatsu *et al*, 2016) were transduced with lentiviruses expressing MS2BP-NLS-GFP and double positive cells were sorted by fluorescence activated cell sorting (FACS). The resulting cell population was maintained without further selection. For control experiments using cytosolic MS2BP (Fig EV2B) stable mCherry-TAF-Iβ cells were transfected with an expression vector for MS2BP-NES-GFP. For RNAscope and immunofluorescene (IF) analysis H1299 cells were grown on 15 mm glass coverslips. Subsequently, cells were infected with 3,000 physical particles/cell of replicative E3-deleted HAd-C5dE3 virus or E1/E3-deleted HAd-C5dE3 GFP expressing control virus, washed once with PBS and fixed at different time points with paraformaldehyde 4%/PBS for 10 min at room temperature. This MOI results in less than 100 virus particles taken up into the cell to not saturate the system but to have a high enough amount of virus for the purpose of this study. Fixed cells were washed 3 times with PBS and slowly dehydrated in 50%/70%/100% ethanol for 5 min each and stored at −20°C in 100% ethanol. For the RNAscope protocol cells were rehydrated for 2 min in 70%/50% ethanol and equilibrated for 10 min in 1× PBS.

Preparation of stable cell lines expression FLAG tagged H3.1 and H3.3

U2OS were seeded in 6-well plates (250,000 cells/well) and transfected with 2 µg of pcDNA3-Histone H3.1-FLAG or pcDNA3-Histone H3.3-FLAG (kindly provided by K. Nagata, Tsukuba University, Japan) using Lipofectamine 2000 (Invitrogen). Two days after transfection, cells were selected with 3 mg/ml G418 for 13 days. Resistant cells were cloned by limiting dilution and the uniform expression of the tagged histones was monitored by immunofluorescence and western-blot on individual clones.

RNAscope

RNAscope experiment were performed using reagents and protocols from the RNAscope® Multiplex Fluorescent Assay essentially as indicated by the supplier (ACD Bio). Briefly, cells were incubated with Protease III (freshly diluted 1:30 with PBS) for 15 min at room temperature (RT) and washed 3 times with PBS. Hybridization with the E1A RNAscope probe was performed for 2 h at 40° in a humidified chamber. Hybridized cells were washed twice for 2 min and hybridized with signal amplifier probes (final amplifier Amp 4 AltB-FL) following the manufacturer's instructions. Hybridized cells were washed twice in PBS followed by incubation in IF buffer (PBS/10% FCS/0.1% Saponin) for 15 min at room temperature to saturate unspecific binding sites followed by IF analysis. Primary antibodies (mouse anti-pVII) were diluted in IF buffer and incubated with cells for 1 h at RT. Antibody was removed by washing in PBS followed by incubation with secondary antibodies (mouse Alexa 647) diluted in IF buffer for 1 h at RT. Cells were washed with PBS, ddH_2O and absolute ethanol, air-dried and mounted in DAKO mixed with DAPI (1:1,000).

Live-cell imaging and image analysis

For live-cell imaging, U2OS cells stably expressing mCherry-TAF-Iβ and MS2BP-NLS-GFP were seeded in ibidi µ-slide $VI^{0.4}$ (Ibidi). The next day, the medium in the ibidi µ-slide was replaced by imaging medium (FluoBrite DMEM supplemented with Prolong antifade reagent, both Life technologies) containing 3,000 physical particles/cell of replicative BxAd5-E1-MS2 for 30 min followed by washing with imaging medium to remove the inoculum. Infection and subsequent cultivation were performed at 37°C. Cells were imaged using a Leica spinning-disk microscopy system (×100 objective) equipped with an environmental chamber at a frame rate of 1 frame per second using MetaMorph software. To generate kymographs, movies were processed using the ImageJ KymographBuilder plugin (https://imagej.net/KymographBuilder_:_Yet_Another_Kymograph_Fiji_plugin) on confocal sections. Line width was adapted to cover all viral genomes throughout the length of the movie and double positive genomes were scored as transcribing genomes. Mounted cells for IF and RNAscope analysis were imaged using a SP8 confocal microscope (Leica) at the Bordeaux Imaging Center (BIC) platform. Stacks (7×) were taken every 0.3 µm using a pinhole of 1 at a pixel size of 85 nm. Quantifications were done using an adapted semi-automated macro in ImageJ (provided upon request). Briefly, channels were split and stacks were Z-projected and the cell periphery was outlined manually. A threshold was applied to every channel and single objects or colocalized objects exceeding a defined area of pixels were counted. Statistical analysis was done using one-way ANOVA multicomparison test.

Virus annotation

The reference genome sequence and gene annotation of Human Adenovirus C serotype 5 (accession: AY339865.1) was downloaded from NCBI. The annotation was adapted to the used Adenovirus strain by excising nucleotides 28,593–30,464 from the full version (modified virus reference genome and gene annotation available at the github project page).

RNA-seq analysis

STAR (Dobin *et al*, 2013) was used to align the RNA-seq reads simultaneously to the human and the adenoviral genome. Therefore, a STAR index was generated based on the hg19 reference, the modified adenoviral genome and gene annotation (see Section Virus annotation for details) and the GENCODE human transcript annotation version 25. Read mapping was conducted using following parameters: --outSAMmultNmax 1 --outMultimapperOrder Random --outFilterType BySJout --outFilterIntronMotifs RemoveNoncanonical --

outSAMstrandField intronMotif --outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --alignIntronMin 20 --alignIntronMax 1000000.

To quantify the RNA abundance for each gene the number of uniquely mapped reads per gene was counted using the featureCounts function of the subread package (Liao *et al*, 2014). The read counts were normalized to transcripts per million (TPMs).

MNase-seq analysis

A bowtie2 index comprised of the human reference genome hg19 and the sequence of the used Adenovirus strain (see Section Virus annotation for details) was built (code available at Schwartz & Längst, 2016) and paired-end reads were mapped to the index using bowtie2-aligner (Langmead & Salzberg, 2012) with following parameters: --very-sensitive-local --no-discordant. Reads were filtered to mapping quality ≥ 30 and only reads still mapped in proper pairs were kept using samtools (Li *et al*, 2009). Reads aligned to the Adenovirus genome were extracted for further analysis. The fragment size distribution was plotted using the CollectInsertSizeMetrics function from the Picard toolkit (https://broadinstitute.github.io/picard; Broad Institute).

pVII positioning analysis

Fragments mapped to the Adenovirus reference with a maximum insert size of 139 bp were selected and used for pVII positioning analysis. pVII positions were called using the DANPOS2 toolkit (Chen *et al*, 2013). First, a pVII occupancy profile for each sample was generated and normalized to 20,000 reads using the dpos function with following parameters: -u 1e-5; -a 1; -z 5; -p 0.001, --extend 35 and --mifrsz 0. Next, the profiles were normalized by a genome wide quantile normalization to the first replicate at 0 hpi of each MNase condition (high/low) using the wiq script. Finally, pVII positions were obtained from the normalized pVII occupancy profiles using the dpos function with following parameters: -jd 50; -q 10; -z 5; -a 1.

Rotational positioning analysis

Forty-nine to fifty-one bp sized fragments mapped to the Adenovirus reference of all high MNase samples were selected and pooled together. Frequencies of A/T or G/C di-nucleotides around the fragment center were calculated using the annotatePeaks.pl script from the HOMER suit (Heinz *et al*, 2010). Noise was filtered using the Fast Fourier Transform build in the filterFFT function of the nucleR Bioconductor package (Flores & Orozco, 2011). The cutoff was set with the option pcKeepComp = 0.2. Finally, to reflect the symmetry of the pVII binding the reverse di-nucleotide frequency was added, and the total frequency was divided by 2. The same procedure was conducted to derive the di-nucleotide periodicity of nucleosomal fragments, except that 147 bp fragments of the first replicate at 0 hpi were selected.

Ad genome accessibility

DNA fragments released under low MNase conditions were used to identify structural changes of Ad chromatin over time. Therefore, fragments mapped to the Ad genome were filtered by a maximum length of 140 bp and replicates of each timepoint were pooled. Next, all timepoints were pooled to call peaks using the DANPOS2 toolkit (Chen *et al*, 2013) with following parameters: dpos -m 1 --extend 35 -c 34,062 -u 0 -z 20 -jd 70 -q 15 -a 1 -e 1. The number of fragments at each timepoint were counted at the determined peaks using featureCounts (Liao *et al*, 2014) with following parameters: --fracOverlap 0.7 -p -B –largestOverlap. Peaks exhibiting in total less than 30 fragment counts were discarded. Fragment counts per peak were normalized to CPMs (counts per million). Changes in chromatin accessibility to 0 hpi at 0.5, 1 and 4 hpi were quantified by an accessibility score. At each timepoint, the accessibility score was obtained for each peak by fitting a linear regression on the logarithmical normalized count frequencies of 0 hpi and the selected timepoint against the logarithmical infection time in min. The slope of the linear fit was used as a measure of dynamic changes of viral genome accessibility (accessibility score). Negative accessibility scores were set to 0, as a further compaction of the viral chromatin is not expected.

Transcription factor motif identification

The Ad DNA sequence from position 19,890 to 19,940 was extracted and uploaded to the PROMO web server (Messeguer *et al*, 2002). Motif search was performed against known human and mouse motifs stored in TRANSFAC data base using the SearchSites tool. The maximum dissimilarity rate was set to 2%. The obtained motifs were further filtered to have at least a recognition sequence of 5 nucleotides and that the corresponding transcription factor is expressed in the infected H1299 cells.

ChIP-seq analysis

Paired-end reads were mapped against viral and human reference genomes and subsequently filtered as described in the MNase-seq analysis section. CPM-normalized coverage tracks were generated using the bamCoverage script of the deepTools package (Ramírez *et al*, 2016). PCA and correlation heatmap were generated using the multiBigwigSummary (50 bp bin size), plotPCA and plotCorrelation deepTools functions.

Sites of nucleosome assembly

The R Bioconductor package nucleR (Flores & Orozco, 2011) was used to call sites of nucleosome assembly onto the Ad genome. First, fragments specific to the Ad genome were loaded into the R environment. Fragments between 137 and 157 bp from time points 1, 2 and 4 hpi of the high MNase condition were extracted for further analysis. The coverage was normalized to reads per million Ad mapped reads and noise was removed using the built-in Fast Fourier Transform (filterFFT with option pcKeepComp = 0.007). Potential nucleosome positions were called based on the normalized profile using the peakDetection function with the option width = 147. Peaks at genomic sites, which exhibit a mean H3.3 ChIP signal (U2OS33) of at least 2 CPM at 4 hpi were used as nucleosome assembly positions.

Data availability

The datasets and computer code produced in this study are available in the following databases: (i) MNase-seq, ChIP-seq and RNA-seq

data: Gene Expression Omnibus GSE136550 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136550); (ii) The code used to analyze the sequencing data: GitHub https://github.com/uschwartz/AdVir. Cell lines and tools are available upon request.

Expanded View for this article is available online.

Acknowledgements

Sequencing was conducted at the NGS Core of the Leibniz Institute for Immunotherapy (LIT, Regensburg, Germany) and at the Genomics Core Unit (GCU, Regensburg, Germany). The study was financed by the BMBF 01DN17003 (to GL), the DFG SFB960 (to GL), the Fondation pour la Recherche Médicale (Equipes FRM 2018 DEQ20180339229 to HW) and by Grants-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (19H04838 to TK). Some of the microscopy was done in the Bordeaux Imaging Center, a service unit of the CNRS-INSERM, and Bordeaux University, a member of the national infrastructure France BioImaging. HW is an INSERM fellow. Open Access funding enabled and organized by Projekt DEAL.

Author contributions

Uwe Schwartz: Data curation; formal analysis; investigation; visualization; methodology; writing – original draft; writing – review and editing. **Tetsuro Komatsu:** Investigation; methodology; writing – review and editing. **Claudia Huber:** Investigation; methodology. **Floriane Lagadec:** Formal analysis; investigation; visualization; methodology. **Conradin Baumgartl:** Investigation; visualization. **Elisabeth Silberhorn:** Investigation; methodology. **Margit Nuetzel:** Methodology. **Fabienne Rayne:** Methodology. **Eugenia Basyuk:** Methodology. **Edouard Bertrand:** Resources; methodology; writing – review and editing. **Michael Rehli:** Investigation; methodology; writing – review and editing. **Harald Wodrich:** Conceptualization; resources; formal analysis; supervision; funding acquisition; investigation; visualization; methodology; writing – original draft; writing – review and editing. **Gernot Laengst:** Conceptualization; resources; supervision; funding acquisition; investigation; methodology; writing – original draft; writing – review and editing.

Disclosure and competing interests statement

The authors declare that they have no conflict of interest.

References

Akusjarvi G (2008) Temporal regulation of adenovirus major late alternative RNA splicing. *Front Biosci* 13: 5006 – 5015

Anderson CW, Young ME, Flint SJ (1989) Characterization of the adenovirus 2 virion protein, mu. *Virology* 172: 506 – 512

Avgousti DC, Herrmann C, Kulej K, Pancholi NJ, Sekulic N, Petrescu J, Molden RC, Blumenthal D, Paris AJ, Reyes ED *et al* (2016) A core viral protein binds host nucleosomes to sequester immune danger signals. *Nature* 535: 173 – 177

Avgousti DC, Della Fera AN, Otter CJ, Herrmann C, Pancholi NJ, Weitzman MD (2017) Adenovirus core protein VII downregulates the DNA damage response on the host genome. *J Virol* 91: e01089-17

Bauer M, Gomez-Gonzalez A, Suomalainen M, Schilling N, Hemmi S, Greber UF (2021) A viral ubiquitination switch attenuates innate immunity and triggers nuclear import of virion DNA and infection. *Sci Adv* 7: eabl7150

Benevento M, Di Palma S, Snijder J, Moyer CL, Reddy VS, Nemerow GR, Heck AJR (2014) Adenovirus composition, proteolysis, and disassembly studied by in-depth qualitative and quantitative proteomics. *J Biol Chem* 289: 11421 – 11430

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM (1998) Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2: 437 – 445

Brison O, Kédinger C, Chambon P (1979) Adenovirus DNA template for late transcription is not a replicative intermediate. *J Virol* 32: 91 – 97

Brown DT, Westphal M, Burlingham BT, Winterhoff U, Doerfler W (1975) Structure and composition of the adenovirus type 2 core. *J Virol* 16: 366 – 387

Cassany A, Ragues J, Guan T, Bégu D, Wodrich H, Kann M, Nemerow GR, Gerace L (2015) Nuclear import of adenovirus DNA involves direct interaction of hexon with an N-terminal domain of the nucleoporin Nup214. *J Virol* 89: 1719 – 1730

Chatterjee PK, Vayda ME, Flint SJ (1985) Interactions among the three adenovirus core proteins. *J Virol* 55: 379 – 386

Chatterjee PK, Vayda ME, Flint SJ (1986) Adenoviral protein VII packages intracellular viral DNA throughout the early phase of infection. *EMBO J* 5: 1633 – 1644

Chen J, Morral N, Engel DA (2007) Transcription releases protein VII from adenovirus chromatin. *Virology* 369: 411 – 422

Chen K, Xi Y, Pan X, Li Z, Kaestner K, Tyler J, Dent S, He X, Li W (2013) DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res* 23: 341 – 351

Chereji RV, Ocampo J, Clark DJ (2017) MNase-sensitive complexes in yeast: nucleosomes and non-histone barriers. *Mol Cell* 65: 565 – 577

Chereji RV, Bryson TD, Henikoff S (2019) Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biol* 20: 198

Corden J, Engelking HM, Pearson GD (1976) Chromatin-like organization of the adenovirus chromosome. *Proc Natl Acad Sci USA* 73: 401 – 404

Crisostomo L, Soriano AM, Mendez M, Graves D, Pelka P (2019) Temporal dynamics of adenovirus 5 gene expression in normal human cells. *PloS One* 14: e0211192

Daniell E, Groff DE, Fedor MJ (1981) Adenovirus chromatin structure at different stages of infection. *Mol Cell Biol* 1: 1094 – 1105

Déry CV, de Murcia G, Lamarre D, Morin N, Poirier GG, Weber J (1986) Possible role of ADP-ribosylation of adenovirus core proteins in virus infection. *Virus Res* 4: 313 – 329

Diermeier S, Kolovos P, Heizinger L, Schwartz U, Georgomanolis T, Zirkel A, Wedemann G, Grosveld F, Knoch TA, Merkl R *et al* (2014) TNFα signalling primes chromatin for NF-κB binding and induces rapid and widespread nucleosome repositioning. *Genome Biol* 15: 536

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15 – 21

Everitt E, Lutter L, Philipson L (1975) Structural proteins of adenoviruses. XII. Location and neighbor relationship among proteins of adenovirion type 2 as revealed by enzymatic lodination, immunoprecipitation and chemical cross-linking. *Virology* 67: 197 – 208

Flint SJ (1982) Expression of adenoviral genetic information in productively infected cells. *Biochim Biophys Acta* 651: 175 – 208

Flores O, Orozco M (2011) nucleR: a package for non-parametric nucleosome positioning. *Bioinformatics* 27: 2149 – 2150

Genoveso MJ, Hisaoka M, Komatsu T, Wodrich H, Nagata K, Okuwaki M (2020) Formation of adenovirus DNA replication compartments and viral DNA accumulation sites by host chromatin regulatory proteins including NPM1. *FEBS J* 287: 205 – 217

Giberson AN, Davidson AR, Parks RJ (2011) Chromatin structure of adenovirus DNA throughout infection. *Nucleic Acids Res* 40: 2369 – 2376

Giberson AN, Saha B, Campbell K, Christou C, Poulin KL, Parks RJ (2018) Human adenoviral DNA association with nucleosomes containing histone variant H3.3 during the early phase of infection is not dependent on viral transcription or replication. *Biochem Cell Biol* 96: 797 – 807

Glenn GM, Ricciardi RP (1988) Detailed kinetics of adenovirus type-5 steady-state transcripts during early infection. *Virus Res* 9: 73 – 91

Greber UF, Willetts M, Webster P, Helenius A (1993) Stepwise dismantling of adenovirus 2 during entry into cells. *Cell* 75: 477 – 486

Hammarskjöld ML, Winberg G (1980) Encapsidation of adenovirus 16 DNA is directed by a small DNA sequence at the left end of the genome. *Cell* 20: 787 – 795

Haruki H, Gyurcsik B, Okuwaki M, Nagata K (2003) Ternary complex formation between DNA-adenovirus core protein VII and TAF-Ibeta/SET, an acidic molecular chaperone. *FEBS Lett* 555: 521 – 527

Haruki H, Okuwaki M, Miyagishi M, Taira K, Nagata K (2006) Involvement of template-activating factor I/SET in transcription of adenovirus early genes as a positive-acting factor. *J Virol* 80: 794 – 801

Hearing P, Samulski RJ, Wishart WL, Shenk T (1987) Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type 5 chromosome. *J Virol* 61: 2555 – 2558

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38: 576 – 589

Hu M, Depledge DP, Flores Cortes E, Breuer J, Schang LM (2019) Chromatin dynamics and the transcriptional competence of HSV-1 genomes during lytic infections. *PLoS Pathog* 15: e1008076

Inturi R, Mun K, Singethan K, Schreiner S, Punga T (2017) Human adenovirus infection causes cellular E3 ubiquitin ligase MKRN1 degradation involving the viral core protein pVII. *J Virol* 92: e01154-17

Jiang C, Pugh BF (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* 10: 161 – 172

Johnson JS, Osheim YN, Xue Y, Emanuel MR, Lewis PW, Bankovich A, Beyer AL, Engel DA (2004) Adenovirus protein VII condenses DNA, represses transcription, and associates with transcriptional activator E1A. *J Virol* 78: 6459 – 6468

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J *et al* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458: 362 – 366

Keller M, Tagawa T, Preuss M, Miller AD (2002) Biophysical characterization of the DNA binding and condensing properties of adenoviral core peptide mu. *Biochemistry* 41: 652 – 659

Kent NA, Adams S, Moorhouse A, Paszkiewicz K (2011) Chromatin particle spectrum analysis: a method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. *Nucleic Acids Res* 39: e26

Komatsu T, Nagata K (2012) Replication-uncoupled histone deposition during adenovirus DNA replication. *J Virol* 86: 6701 – 6711

Komatsu T, Haruki H, Nagata K (2011) Cellular and viral chromatin proteins are positive factors in the regulation of adenovirus gene expression. *Nucleic Acids Res* 39: 889 – 901

Komatsu T, Dacheux D, Kreppel F, Nagata K, Wodrich H (2015) A method for visualization of incoming adenovirus chromatin complexes in fixed and living cells. *PloS One* 10: e0137102

Komatsu T, Nagata K, Wodrich H (2016) An adenovirus DNA replication factor, but not incoming genome complexes, targets PML nuclear bodies. *J Virol* 90: 1657 – 1667

Kubik S, Bruzzone MJ, Jacquet P, Falcone JL, Rougemont J, Shore D (2015) Nucleosome stability distinguishes two different promoter types at all protein-coding genes in yeast. *Mol Cell* 60: 422 – 434

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357 – 359

Längst G, Schätz T, Langowski J, Grummt I (1997) Structural analysis of mouse rDNA: coincidence between nuclease hypersensitive sites, DNA curvature and regulatory elements in the intergenic spacer. *Nucleic Acids Res* 25: 511 – 517

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078 – 2079

Liao Y, Smyth GK, Shi W (2014) FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923 – 930

Lischwe MA, Sung MT (1977) A histone-like protein from adenovirus chromatin. *Nature* 267: 552 – 554

Martinez R, Schellenberger P, Vasishtan D, Aknin C, Austin S, Dacheux D, Rayne F, Siebert A, Ruzsics Z, Gruenewald K *et al* (2015) The amphipathic helix of adenovirus capsid protein VI contributes to penton release and postentry sorting. *J Virol* 89: 2121 – 2135

Matsumoto K, Okuwaki M, Kawase H, Handa H, Hanaoka F, Nagata K (1995) Stimulation of DNA transcription by the replication factor from the adenovirus genome in a chromatin-like structure. *J Biol Chem* 270: 9645 – 9650

Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM (2002) PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* 18: 333 – 334

Mieczkowski J, Cook A, Bowman SK, Mueller B, Alver BH, Kundu S, Deaton AM, Urban JA, Larschan E, Park PJ *et al* (2016) MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nat Commun* 7: 11485

Minderjahn J, Schmidt A, Fuchs A, Schill R, Raithel J, Babina M, Schmidl C, Gebhard C, Schmidhofer S, Mendes K *et al* (2020) Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat Commun* 11: 402

Mirza MA, Weber J (1982) Structure of adenovirus chromatin. *Biochim Biophys Acta* 696: 76 – 86

Mittereder N, March KL, Trapnell BC (1996) Evaluation of the concentration and bioactivity of adenovirus vectors for gene therapy. *J Virol* 70: 7498 – 7509

Mueller B, Mieczkowski J, Kundu S, Wang P, Sadreyev R, Tolstorukov MY, Kingston RE (2017) Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction. *Genes Dev* 31: 451 – 462

Mun K, Punga T (2018) Cellular zinc finger protein 622 hinders human adenovirus lytic growth and limits binding of the viral pVII protein to virus DNA. *J Virol* 93: e01628-18

Nevins JR, Ginsberg HS, Blanchard JM, Wilson MC, Darnell JE (1979) Regulation of the primary expression of the early adenovirus transcription units. *J Virol* 32: 727 – 733

Noll M, Thomas JO, Kornberg RD (1975) Preparation of native chromatin and damage caused by shearing. *Science* 187: 1203 – 1206

Ostapchuk P, Suomalainen M, Zheng Y, Boucke K, Greber UF, Hearing P (2017) The adenovirus major core protein VII is dispensable for virion assembly but is essential for lytic infection. *PLoS Pathog* 13: e1006455

Pheasant K, Möller-Levet CS, Jones J, Depledge D, Breuer J, Elliott G (2018) Nuclear-cytoplasmic compartmentalization of the herpes simplex virus 1 infected cell transcriptome is co-ordinated by the viral endoribonuclease vhs and cofactors to facilitate the translation of late proteins. *PLoS Pathog* 14: e1007331

Pied N, Wodrich H (2019) Imaging the adenovirus infection cycle. *FEBS Lett* 593: 3419 – 3448

Puntener D, Engelke MF, Ruzsics Z, Strunze S, Wilhelm C, Greber UF (2011) Stepwise loss of fluorescent core protein V from human adenovirus during entry into cells. *J Virol* 85: 481 – 496

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

Ross PJ, Kennedy MA, Christou C, Risco Quiroz M, Poulin KL, Parks RJ (2011) Assembly of helper-dependent adenovirus DNA into chromatin promotes efficient gene expression. *J Virol* 85: 3950 – 3958

Ruzsics Z, Lemnitzer F, Thirion C (2014) Engineering adenovirus genome by bacterial artificial chromosome (BAC) technology. *Methods Mol Biol* 1089: 143 – 158

Samad MA, Komatsu T, Okuwaki M, Nagata K (2012) B23/nucleophosmin is involved in regulation of adenovirus chromatin structure at late infection stages, but not in virus replication and transcription. *J Gen Virol* 93: 1328 – 1338

Sato K, Hosokawa K (1984) Analysis of the interaction between DNA and major core protein in adenovirus chromatin by circular dichroism and ultraviolet light induced cross-linking. *J Biochem* 95: 1031 – 1039

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132: 887 – 898

Schwartz U, Längst G (2016) Bioinformatic analysis of ChIP-seq data on the repetitive ribosomal RNA gene. *Methods Mol Biol* 1455: 225 – 230

Schwartz U, Németh A, Diermeier S, Exler JH, Hansch S, Maldonado R, Heizinger L, Merkl R, Längst G (2018) Characterizing the nuclease accessibility of DNA in human cells to map higher order structures of chromatin. *Nucleic Acids Res* 47: 1239 – 1254

Sekine E, Schmidt N, Gaboriau D, O'Hare P (2017) Spatiotemporal dynamics of HSV genome nuclear entry and compaction state transitions using bioorthogonal chemistry and super-resolution microscopy. *PLoS Pathog* 13: e1006721

Strunze S, Engelke MF, Wang I-H, Puntener D, Boucke K, Schleich S, Way M, Schoenenberger P, Burckhardt CJ, Greber UF (2011) Kinesin-1-mediated capsid disassembly and disruption of the nuclear pore complex promote virus infection. *Cell Host Microbe* 10: 210 – 223

Suomalainen M, Prasad V, Kannan A, Greber UF (2020) Cell-to-cell and genome-to-genome variability of adenovirus transcription tuned by the cell cycle. *J Cell Sci* 134: jcs252544

Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A (2011) Determinants of nucleosome organization in primary human cells. *Nature* 474: 516 – 520

Vayda ME, Rogers AE, Flint SJ (1983) The structure of nucleoprotein cores released from adenovirions. *Nucleic Acids Res* 11: 441 – 460

Wickham TJ, Mathias P, Cheresh DA, Nemerow GR (1993) Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. *Cell* 73: 309 – 319

Widom J (2001) Role of DNA sequence in nucleosome stability and dynamics. *Q Rev Biophys* 34: 269 – 324

Wiethoff CM, Wodrich H, Gerace L, Nemerow GR (2005) Adenovirus protein VI mediates membrane disruption following capsid disassembly. *J Virol* 79: 1992 – 2000

Xue Y, Johnson JS, Ornelles DA, Lieberman J, Engel DA (2005) Adenovirus protein VII functions throughout early phase and interacts with cellular proteins SET and pp32. *J Virol* 79: 2474 – 2483