

Changes in adenoviral chromatin organization precede early gene activation upon infection

Uwe Schwartz, Tetsuro Komatsu, Claudia Huber, Floriane Lagadec, Conradin Baumgartl, Elisabeth Silberhorn, Margit Nuetzel, Fabienne Rayne, Eugenia Basyuk, Edouard Bertrand, Michael Rehli, Harald Wodrich, and Gernot Längst

DOI: [10.15252/embj.2023114162](https://doi.org/10.15252/embj.2023114162)

Corresponding author(s): Gernot Längst (gernot.laengst@ur.de) , Harald Wodrich (harald.wodrich@u-bordeaux.fr)

Review Timeline:

Transfer from Review Commons:	31st Mar 23
Editorial Decision:	21st Apr 23
Revision Received:	10th Jul 23
Accepted:	4th Aug 23

Review
COMMONS

Editor: Ieva Gailite

Transaction Report: This manuscript was transferred to The EMBO JOURNAL following peer review at Review Commons.

Review #1

1. Evidence, reproducibility and clarity:

Evidence, reproducibility and clarity (Required)

The authors have addressed the nucleoprotein structure of human adenovirus during the very early stages of infection, and its relationship to onset of expression of viral genes, using a combination of RNA-seq, MNase-seq, ChIP-seq and single genome imaging. They show that in the virion and the newly-infecting DNA, protein VII is precisely positioned at specific sites on the viral DNA, with greater accessibility at early gene promoters compared to other regions. Nucleosomes containing H3.3 replace specific protein VII at distinct positions at the transcription start sites of genes, which are then acetylated. Association with histones and nucleosomes occurs prior to transcription. These studies confirm and greatly expand on results already in the literature, and also elucidate a novel role for protein VII in orchestrating positioning of nucleosomes prior to initiation of transcription.

The authors provide excellent data in support of their conclusions and, in many instances, use alternative experiments (i.e. two different approaches) to support their claims. The details of methods are adequate (with small exceptions outlined below) and statistical methods appropriate.

****Minor comments:****

1. Line 561 "Protein VII molecules were exchanged for positioned nucleosomes at the +1 site of actively transcribed genes". This statement seems to suggest that the +1 position almost acts as a nucleating site, where replacement of a single, specific protein VII molecule at +1 is an initiating event, which then spreads from that site and into the rest of the gene. Data shown in Figure 6G and 6H shows that H3.3 appears to be found equally along the full length of E1A as early as 1 hr post infection (with no real "enhancement" at the +1 position), and that the overall levels simply increase over the next 4 hrs.
2. Curiously, the authors chose not to use a wildtype virus for their studies - the virus contains a deletion in the E3 region. For clarity, I suggest that the authors should preferentially use an alternative designation for their virus rather than HAd-C5. Perhaps HAd-C5delE3 to differentiate this work from studies that truly use wildtype virus.
3. The obvious limitation of the studies using the fluorescent TAF1-beta to label Ad genomes is that as protein VII is replaced by nucleosomes, the genomes would have

declining detection by this method. Genomes devoid of protein VII would be "invisible".

4. Line 275 "Interestingly, a central region of the viral genome (Late3) and a region between the E3 and E4 genes exhibited almost no peaks" for protein VII. The virus utilized in this study lacked at least part of the E3 region. Did this deletion "cause" this region to be devoid of protein VII? Is the same absence of protein VII peaks observed in a fully wildtype virus? Also, can the authors provide any speculation as to why the Late3 region also lacks protein VII?

5. Line 569 "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." This argument may make sense in the context of a high MOI infection, but would certainly limit virus function during normal, pathogenic infection where the MOI is likely extremely low. Essentially, the authors data predicts that 80% of normal, low MOI infections don't progress to gene expression (at least during the first 4 hrs analyzed in this study).

6. Line 576 "This observation is in agreement with recent pVII-ChIP experiments showing transcription and replication independent pVII removal in early infection (Giberson et al., 2018; Komatsu and Nagata, 2012; Komatsu et al., 2011)." The authors can also state that histone and nucleosome deposition is also independent of transcription and replication, as has been alluded to in the same cited studies but proven more directly in this study.

7. Line 672 - the authors should be more definitive in the MOI that are used in all of their experiments. Line 672 states that an MOI of 3000 physical particles are applied per cell. There can be great variation between cell lines in how much virus binds to (and enters) a cell based on the surface levels of Ad receptors on different cell types. However, in general, 3000 is very high. Work by Wang et al. (PMID:24139403) showed that at an MOI of 200 or below most Ad will traffic correctly to the nucleus, whereas at an MOI above 200 there is a significant defect in Ad trafficking within the cell. How is this expected to affect all of the results in this study?

8. Figure 5 is of low resolution and was difficult to read.

9. Figure S3 is missing a box from the top set of images indicating the region that is expanded in the detail picture.

10. While I realize it is supplemental data, the difference in quality between the agarose gels shown in Figure S4A and S5A is shocking.

11. Figure S7 is of low resolution.

2. Significance:

Significance (Required)

At least in the field of adenovirus research, this is a very important study. There has been considerable debate in the field regarding the timing and degree of protein VII removal and histone deposition, and the necessity of active transcription for these two events. The data provided in this manuscript clearly shows that some protein VII is removed from early active genes and replaced by nucleosomes, and that these events occur prior to initiation of transcription.

The authors speculate that the specific placement of protein VII, a protamine-like protein, on the Ad genome prescribes where nucleosomes are placed. This finding should be of interest to a broad general audience, as it provides novel information on chromatin assembly within mammalian cell.

Key words for this reviewer: adenovirus research, HAdV nucleoprotein structure

3. How much time do you estimate the authors will need to complete the suggested revisions:

Estimated time to Complete Revisions (Required)

(Decision Recommendation)

Less than 1 month

4. *Review Commons* values the work of reviewers and encourages them to get credit for their work. Select 'Yes' below to register your reviewing activity at [Web of Science Reviewer Recognition Service](#) (formerly Publons); note that the content of your review will not be visible on Web of Science.

Web of Science Reviewer Recognition

No

Review #2

1. Evidence, reproducibility and clarity:

Evidence, reproducibility and clarity (Required)

The submitted manuscript presents a detailed and comprehensive analysis of the adenoviral nucleoprotein complexes as infection progresses, starting with the "adenosome" assembled with pVII which are then progressively replaced with H3.3-containing nucleosomes as the infection progresses. The submission presents a combination of in situ and populational analyses of the viral DNA accessibility and complexes through infection. In brief, the infecting viral genomes are assembled in some 250 adenosomes with pVII, which become progressively replaced as infection progresses with nucleosomes containing H3.3 and acetylated H3K17, starting at the active promoters of the E genes. Chromatin remodeling precedes transcription, and the accessibility differs for genes of different kinetic classes at different times after infection, although there is no correlation between accessibility and H3.3 or acetylation content. Only about 20% of the genomes become transcriptionally active, though, which somewhat complicates the analyses of the populational studies of accessibility and occupancy. Overall, the study is well conceived, performed and presented. A few issues that deserve further analyses and discussion, as described below.

****Major issues.****

As figure 2 nicely shows, only about 20% of the intranuclear genomes become transcriptionally active. However, MNase and ChIP analyses cannot differentiate these genomes from the 80% that are transcriptionally inactive. The interpretation of the positioning of pVII (figure 4) or the changes in compaction of the adenoviral chromatin at different loci (figure 5) does not appear to consider this heterogeneity other than for a brief comment about the stringent MNase digestion in page 11. The authors favor a model in which the changes in compaction shown in figure 5, at mild MNase digestions, directly correlate with transcription of the respective genes. This could well be correct, and in fact the correlation may be underestimated as 80% of the genomes may not undergo any changes, but it may also be incorrect. The analyses presented cannot differentiate whether the changes in chromatin compaction occur in only a subset of genomes or in all the genomes, regardless of whether they are transcribed or not, or even only in the non-transcribed genomes (which appears extremely unlikely). This intrinsic limitation to the methods used (and I know of no

better alternative) should be acknowledged and discussed for the benefit of the reader. This limitation also impacts the analyses of the lack of correlation between H3.3 and acetylated H3K27 occupancy and compaction.

Perhaps out of necessity to reach the required sensitivity, a high multiplicity of infection was used (although the actual moi is not stated, there are about 25-30 pVII foci/ per nuclei). The presentation, analyses and discussion of the results should emphasize this context. For example, one would presume that at low moi, when only one genome enters each cell, the percentage of transcriptionally active genomes in a given cell will be either 0 or 100%, but the "system" becomes saturated as more and more genomes enter the nucleus at higher moi resulting in only a subset of them being transcriptionally active. Along this line of reasoning, it is intriguing that the percentage of genomes estimated to be in nucleosomes at 4 hpi (14%) approaches the percentage of transcribed genomes.

The changes in chromatin compaction presented in figure 5 are in some respect puzzling. The compaction of most of the late genes increases as infection progresses, at least for the first four hours, as the authors discuss. However, the L genes appear to be at least as accessible as the E ones at the early times, when only the E are transcribed to high levels. This appears counterintuitive, and may not be consistent with the main conclusion that increase accessibility to a given gene directly correlates to its transcriptional activity level. The data presented in Figure 5C deserves a more nuanced analysis and discussion, parsing out the changes in accessibility to each given gene at different times from the different accessibility to the different genes at any given time. The later does not appear to support the main conclusion reached by the authors that accessibility to each individual gene correlates with its transcriptional level.

****Minor comments****

The authors may wish to highlight in the discussion that the analyses are so far limited to a single adenovirus.

The y-axes in the transcriptome figures (figure 1 B, S2) could be presented in Log(2) scale, such that transcript levels at all times can be appreciated in the same graph (the earlier times are just not visible in a linear scale)

The (lack of) phenotype of the 24xMS2 binding site recombinant adenovirus used should be shown.

The kymograph analyses presented in figure 3B appear to show that there are some sites of transcript accumulation sites which do not harbor viral genomes (i.e., green only tracks). Moreover, the interpretation of the TAF1beta-mCherry signal is complicated by the (fully expected) significant "background" signal. Although these results are consistent with those obtained by RNAscope/pVII staining, there appears to be intrinsic limitations to the system, which preclude reaching strong conclusions from it. These confirmatory analyses should probably be moved to the supplementary information section and removed from the main text and figures. The longer evaluation data mentioned as not shown in page 8 is critical to the conclusions and should be shown.

Although the plot of cleavage frequency presented in figure S5 is clear, it would be beneficial to the reader if the actual peaks were also presented to compare their distribution (if any) in gDNA and virus particle.

The mRNA analyses of selected transcription factors provides little information, as there is no context, there is variability between experiments, and in most cases the changes appear modest. As these results are not critical to the conclusions or analyses, perhaps the authors may wish to remove them from the manuscript. Alternatively, more in-depth analyses would be required.

It is unclear why the even distribution of H3.1-flag signal across the genome is considered indicative of no specific recruitment. The results presented are equally consistent with equal incorporation across the genome. Perhaps the authors have some additional information, such as an irrelevant antibody, input DNA, or the like, to support the conclusion. If so, that evidence should be presented and discussed. If not, the interpretation should be revisited. As an added complexity, endogenous H3.1 is normally expressed during S-phase. It is possible that Adenovirus infection may induce higher levels of expression of (untagged)endogenous H3.1, which would outcompete the tagged ectopically expressed histone. These analyses deserve a more nuanced and in-depth analysis.

It is highly unlikely that the somewhat higher H3K27ac signal observed in the H3.3 than in the H3.1 expressing cells may result from higher H3.3. occupancy in the viral genome as speculated in page 13. The total levels of H3.3. are unlikely to increase by the ectopically expressed one, and even if they did it is not likely that the occupancy of the viral genome would be limited by the levels of H3.3. This speculation should be removed.

Materials and methods are too concise. A longer more detailed version, as supplementary information, would be highly desirable.

2. Significance:

Significance (Required)

The major strengths of this manuscript lie on its comprehensiveness, using several in situ and populational approaches to address biologically critical questions regarding the regulation of viral replication by chromatin and epigenetics. Experiments appear very well designed and performed and are mostly clearly presented. The interpretation analyses and discussion of the results may benefit from a more nuanced analysis of the issues posed by the existence of different populations of viral genomes in the cells infected at high moi and the accessibility across different genes at any given time versus the levels of transcription of the different genes, which appears not to be fully consistent with one of the main conclusions reached.

This study makes a very significant contribution, describing the dynamic changes in the adenoviral nucleoprotein complexes at the early times of infection and providing a full description of both the adenosomes and the nucleosomes in more and less transcribed loci. The results are properly analyzed in context of what is known about the regulation of viral gene transcription by chromatin dynamics in other systems, including similarities and differences. This study is likely to be of high interest to a wide audience, ranging from virologist to epigeneticists, to those working in gene therapy and vectored vaccines.

3. How much time do you estimate the authors will need to complete the suggested revisions:

Estimated time to Complete Revisions (Required)

(Decision Recommendation)

Less than 1 month

4. *Review Commons* values the work of reviewers and encourages them to get credit for their work. Select 'Yes' below to register your reviewing activity at [Web of Science Reviewer Recognition Service](#) (formerly Publons); note that the content of your review will not be visible on Web of Science.

Web of Science Reviewer Recognition

Yes

Review #3

1. Evidence, reproducibility and clarity:

Evidence, reproducibility and clarity (Required)

The manuscript "Adenoviral chromatin organization primes for early gene activation" combines RNA-seq, MNase-seq, ChIP-seq, and single genome and transcript imaging (immunofluorescence, RNA-scope, and live cell techniques) during early Adenovirus infection in vitro to characterise the spatiotemporal dynamics of viral chromatin organisation and association with gene transcription. The manuscript is an interesting read and the authors have combined multiple complimentary techniques to make a substantial contribution to understanding the early events occurring after nuclear import of viral genomes.

Adenoviruses are important causes of human and animal pathology, are a useful model of non-integrating extra-chromosomal DNA virus infection in mammalian cells, and are useful vectors for vaccination and the discoveries may influence gene therapy DNA vector design. The chromatin organisation in adenovirus infection is distinct from other DNA viruses, and is relatively poorly understood compared to, for example, SV40 or herpesviruses. The manuscript describes an early transition from purely viral chromatin with Adenovirus protein pVII packaging the virus in virions, to a viral-human hybrid chromatin pattern with apparently strategically positioned H3.3 nucleosomes and viral pVII "Adenosomes" in the early hours after nuclear import of the viral genome. The data shows that packaged Adenoviruses are in a transcriptionally accessible form and gene expression occurs rapidly after infection, the combination of the MNase-seq data with ChIP-seq data is particularly interesting demonstrating an average ~238 adenosomes positioned by specific DNA code protecting 60-70bp of DNA, and that the genome is accessible at loci that also decondense on infection, with adenosomes being replaced by cellular H3.3 containing nucleosomes at distinct sites. Particularly they show that +1 H3K27 acetylated nucleosomes are acquired at the TSS of key early genes.

The authors argue that their spatiotemporal data imply that this chromatin transition

"primes" for early gene transcription. The manuscript is well written, uncovers important viral chromatin biology by combining multiple experimental techniques, and the data is generally very clearly presented. A few comments follow.

****Major concerns:****

- Abstract and Title:

- the abstract and title suggest that because the chromatin changes are observed coincidentally or before transcriptional changes, and that this means that these chromatin changes "prime" (title) or are "required" and play a "central role" (abstract) in early gene expression. The temporal relationship would be consistent with chromatin changes being required for transcriptional changes, but do not imply necessity. Experiments to demonstrate the necessity of these changes for early gene transcription are lacking, and I recommend amending the text or additional experiments to provide this evidence directly.

- Results:

- The IF data in Fig S1 is convincing, showing viral particles are accessible quickly in the nucleus. Although no statistics are provided for S1B and C, pVII foci appear at 0.5hpi and appear to mostly accumulate between 0.5hpi and 1hpi with further import between 1hpi and 4hpi. Can the authors be sure that a single pVII IF focus represents a single genome? If genomes tend to aggregate as they accumulate the number of foci per nucleus may not increase linearly with the number of genomes imported. Have the authors considered analysing the intensity of the individual pVII foci over the time points? A related question is whether the authors assume that all packaged virions contain intact complete viral genomes? Many viruses comprise some mixture of complete and incomplete packaged genomes, and the subsequent analyses determine the proportion of transcriptionally active copies with RNA-Scope to a single transcript E1A which lies at one end of the viral genome. Please comment explicitly on whether this is assumed and whether this assumption is realistic in light of known Adenovirus biology.

- The RNA-Seq data in Fig 1 and Fig S2 and Table S1 demonstrates transcription of early genes is barely observable at 1hpi but is observable by 2hpi and is clearly much increased by 4hpi. Fig 2C, visualising pVII foci directly within single cells, suggests that approximately 80% of foci are observed by 1hpi and a further 20% between 1hpi and 2hpi and little thereafter. These data convincingly demonstrate that nuclear import is rapid, typically occurring in the first hour. The E1A RNA-Scope data in figure 2, visualising individual mRNA transcripts of E1A, is more sensitive than the bulk RNA-Seq, and shows transcripts at 1hpi with clearly discernible transcription by 2hpi (2A&D) which suggests that transcription occurs early, by 2hpi. Thus transcription lags nuclear genome import by approximately one hour by these methods. However, the conclusions of the subsequent analyses depend on the chromatin changes clearly

preceding, rather than being approximately coincident with transcription, therefore transcription being evident by 2hpi is relevant as figure 6A and D suggest that the chromatin remodelling is subtle before 2hpi on the bulk sequencing analyses. The authors should comment on this given the importance to their argument.

- The validation of the E1A probe specificity in Fig 2B looks convincing, but there are no data presented for multiple cells to reassure that this image is representative. The equivalent figure for 2D for the Ad5-GFP control would address this.

- Figure 2E is presented as a colocalization analysis but appears to be a ratio of mRNA foci to pVII foci per cell. If this is an incorrect interpretation then some clarification in the figure legend would be helpful. If this interpretation of these data is correct, then it is not truly a colocalization analysis, as a single genome may give rise to multiple transcripts and so a ratio <1 could be expected even if all pVII foci were transcribing genomes. In addition, the text suggests that because there is no statistically significant difference between 2hpi and 4hpi a plateau has been reached. However, the difference between 1h and 2h is barely significant ($p=0.04$) and the mean is increased between 2h and 4h, albeit non-significantly, so the plateau is not convincingly demonstrated. If the authors wish to perform a colocalization analysis rather than a ratio, they might assign each transcript to the nearest pVII IF focus within the nucleus and count the proportion of pVII foci with any transcripts assigned over time. Alternatively, they can amend the description of this analysis in the figure and text.

- The live cell imaging experiments are elegant and convincing, but the agreement in Fig 3D of the % colocalization in MS2-BP data with the RNA-scope data is potentially misleading for the reasons outlined in the prior comment. Is the data in Fig 2E the same as the data in the right hand panel of Fig 3D. If so please comment on the n discrepancy ($n=30$ in 2E vs $n=22$ in 3D). The observation that 20% of genomes are transcriptionally active, via bursting or otherwise, is interesting, and would be consistent with the Suomalainen et al reference. The authors discuss two hypotheses to explain these findings: transcriptional bursting or a subset ~20% of genomes being transcriptionally active. This is an interesting and begs the question as to why this may occur. Assuming all imported genomes are intact (previous comment), it appears from the presented images that the foci at the radial periphery of the nucleus may be more frequently transcriptionally active, despite the nuclear periphery being enriched for heterochromatin. The authors might consider analysing the radial position of their TAF1B-mCherry genomes (active and inactive) as this might support position effect variegation rather than bursting as an explanation and they appear to already have the data to perform such analyses.

- In the presented images (Fig 3A and Fig S3) it appears a higher proportion of genomes than 20% appear to be transcriptionally active, particularly in the low MOI experiment. The authors may wish to comment on this and quantify whether the proportion of transcribing genomes was affected by the input MOI.

- Fig 4C suggests that there is a large GC preference (or bias) in the pVII occupied regions. The authors may wish to comment on this and present a track with Adenovirus GC composition in Fig 4D.
- Figure 6 presents convincing data showing H3.3. nucleosome positioning and acetylation at E1A and the data is nicely presented showing these changes occur early being observable by 2 and 4 hpi. Again, these changes are not convincingly prior to early gene activation but are certainly occurring early, and may occur prior to early gene activation at the level of individual foci, however, this is not demonstrated definitively.

****Minor comments:****

- Introduction:
- Paragraph 1 - Introduction for DNA viruses in general, but the authors appear to be talking about Adenoviruses specifically, "little is known about the structural organization of the genome" and "nuclear viral genomes could undergo different parallel fates", arguably these statements are not accurate for other DNA viruses (e.g. Epstein Barr Virus) suggest amending the wording for clarity.
- paragraph 2 - Why do the authors say that Adenoviruses are prototypic DNA viruses?
- Paragraph 3 - A recent study is referenced but multiple references are given.
- "Protein VII stays associated with the viral genome imported to the nucleus, while pV dissociates from the viral DNA following ubiquitylation (Puntener et al., 2011). The fate of the μ -peptide is not known". - The reference suggests that pV dissociates on entry to the cytoplasm and during capsid disassembly at the nuclear pore. I find this sentence confusing as it doesn't make it clear that pV is lost before nuclear entry which is important for interpreting the data.
- Results:
- Figure 5 is almost unreadable due to low resolution.
- Reference to Fig 4C in text comes after Fig4D.

2. Significance:

Significance (Required)

The manuscript is well written, uncovers important viral chromatin biology by combining multiple experimental techniques, and the data is generally very clearly presented

3. How much time do you estimate the authors will need to complete the suggested revisions:

Estimated time to Complete Revisions (Required)

(Decision Recommendation)

Between 3 and 6 months

4. *Review Commons* values the work of reviewers and encourages them to get credit for their work. Select 'Yes' below to register your reviewing activity at [Web of Science Reviewer Recognition Service](#) (formerly Publons); note that the content of your review will not be visible on Web of Science.

Web of Science Reviewer Recognition

No

Full Revision



Manuscript number: RC- 2023-01819

Corresponding author(s): Gernot Längst and Harald Wodrich

Full revision of the manuscript

1. General Statements [optional]

This section is optional. Insert here any general statements you wish to make about the goal of the study or about the reviews.

2. Point-by-point description of the revisions

Dear Reviewers, thank you very much for your appreciation of our study and your input. In this point-to-point response, we amended our text marked in blue colour.

Reviewer #1 (Evidence, reproducibility and clarity (Required)):

The authors have addressed the nucleoprotein structure of human adenovirus during the very early stages of infection, and its relationship to onset of expression of viral genes, using a combination of RNA-seq, MNase-seq, ChIP-seq and single genome imaging. They show that in the virion and the newly-infecting DNA, protein VII is precisely positioned at specific sites on the viral DNA, with greater accessibility at early gene promoters compared to other regions. Nucleosomes containing H3.3 replace specific protein VII at distinct positions at the transcription start sites of genes, which are then acetylated. Association with histones and nucleosomes occurs prior to transcription. These studies confirm and greatly expand on results already in the literature, and also elucidate a novel role for protein VII in orchestrating positioning of nucleosomes prior to initiation of transcription.

The authors provide excellent data in support of their conclusions and, in many instances, use alternative experiments (i.e. two different approaches) to support their claims. The details of methods are adequate (with small exceptions outlined below) and statistical methods appropriate.

Minor comments:

1. Line 561 "Protein VII molecules were exchanged for positioned nucleosomes at the +1 site of actively transcribed genes". This statement seems to suggest that the +1 position almost acts as a nucleating site, where replacement of a single, specific protein VII molecule at +1 is an initiating event, which then spreads from that site and into the rest of the gene. Data shown in Figure 6G and 6H shows that H3.3 appears to be found equally along the full length of E1A as early as 1 hr post infection (with no real "enhancement" at the +1 position), and that the overall levels simply increase over the next 4 hrs.

As the reviewer pointed out, the histone ChIP-seq peaks are broader than the +1 nucleosome region, extending into the transcribed regions of the gene. This is expected, as the mean length of

the immunoprecipitated DNA is about 400bp long. Still, ChIP-seq peaks are in proximity to the transcription start site and overlap with the position of the +1 nucleosome. As we do not have the required resolution, we toned down our statement. The text now reads as follows: "Protein VII molecules were exchanged for nucleosomes downstream of the transcription start site, overlapping the +1 nucleosome site, of actively transcribed genes" (line 568 ff).

2. Curiously, the authors chose not to use a wildtype virus for their studies - the virus contains a deletion in the E3 region. For clarity, I suggest that the authors should preferentially use an alternative designation for their virus rather than HAd-C5. Perhaps HAd-C5delE3 to differentiate this work from studies that truly use wildtype virus.

As requested by the reviewer we have updated the nomenclature to HAd-C5delE3 throughout the text and the figures.

3. The obvious limitation of the studies using the fluorescent TAF1-beta to label Ad genomes is that as protein VII is replaced by nucleosomes, the genomes would have declining detection by this method. Genomes devoid of protein VII would be "invisible".

Our MNase data show that within the first 4h only a fraction of pVII is removed from the viral genome e.g. at early genes, while most of the genome remains bound by protein VII. This should provide enough binding sites for TAF1-beta to label Ad genomes without a significant drop in the signal. Furthermore, our recent work (PMID:29997215, Fig. 1D) compared the TAF1-beta labelling system with a second *in vivo* detection system (AnchOR3) that directly labels the viral DNA independently of protein VII in the same cells. This direct comparison of two technically non-related methods to detect individual incoming adenoviral genomes in living cells showed the equivalence of both methods, at least for the first hours of infection showing that partial removal of protein VII does not affect the fluorescent TAF1-beta staining.

4. Line 275 "Interestingly, a central region of the viral genome (Late3) and a region between the E3 and E4 genes exhibited almost no peaks" for protein VII. The virus utilized in this study lacked at least part of the E3 region. Did this deletion "cause" this region to be devoid of protein VII? Is the same absence of protein VII peaks observed in a fully wildtype virus? Also, can the authors provide any speculation as to why the Late3 region also lacks protein VII?

We confirm the reviewer's observation. The region marked as Late3 and the region between E3 and E4 is present in the genome and is, as the reviewer observed, not chromatinized in our analysis. At this point, we can only speculate. We have two not mutually exclusive hypotheses. First, both regions could be involved in the proper packaging of the viral genome into the capsid. Physical constraints during packaging may preclude this region from being packaged into pVII. Second, as we observed that pVII positioning correlates with distinct DNA sequence patterns (revised Fig. 4 D and E, see response to reviewer 3 for details), it might be that the sequence composition at the pVII depleted regions disfavour pVII assembly to keep those regions available for cellular factors that drive processes post genome delivery, such as transcription. Our time-resolved MNase analysis shows that indeed post genome delivery, this site in the Late3 region becomes protected (Fig. 5C), suggesting the binding of one or more cellular factors. As shown in Figure S6 we find conserved binding sites for several transcription factors at this MNase protected site.

Whether the chromatinization devoid regions would shift in position, remain in place or be chromatinized in a wildtype virus has to be addressed in the future and cannot be answered at this point. To address the comment, we have expanded the discussion (line 620 ff)

5. Line 569 "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." This argument may make sense in the context of a high MOI infection, but would certainly limit virus function during normal, pathogenic infection where the MOI is likely extremely low. Essentially, the authors data predicts that 80% of normal, low MOI infections don't progress to gene expression (at least during the first 4 hrs analyzed in this study).

We follow the argument of the reviewer. The high MOI in our study was necessary to perform the combined 'omics' approach to arrive at meaningful data within reasonable sequencing depth. To have equivalence we also used high MOI for the imaging approach. A detailed analysis for the effect of low MOI as well as positioning effects (see reviewer comment below) on transcriptional activation is an important question and will be addressed in future studies that require different techniques in addition. To address this comment, we have updated the discussion to emphasize the importance of MOI and positioning effects (line 587 ff).

6. Line 576 "This observation is in agreement with recent pVII-ChIP experiments showing transcription and replication independent pVII removal in early infection (Giberson et al., 2018; Komatsu and Nagata, 2012; Komatsu et al., 2011)." The authors can also state that histone and nucleosome deposition is also independent of transcription and replication, as has been alluded to in the same cited studies but proven more directly in this study.

We have changed the text accordingly (line 576 and 598).

7. Line 672 - the authors should be more definitive in the MOI that are used in all of their experiments. Line 672 states that an MOI of 3000 physical particles are applied per cell. There can be great variation between cell lines in how much virus binds to (and enters) a cell based on the surface levels of Ad receptors on different cell types. However, in general, 3000 is very high. Work by Wang et al. (PMID:24139403) showed that at an MOI of 200 or below most Ad will traffic correctly to the nucleus, whereas at an MOI above 200 there is a significant defect in Ad trafficking within the cell. How is this expected to affect all of the results in this study?

We agree with this and the other reviewer that this is an important issue. The actual dose of virus that enters a given cell is dependent on the concentration of virus particles in the inoculum and the time and temperature this inoculum is in contact with the cells and the cells respective susceptibility to the virus. We applied an infection dose of 3000 physical particles per cell in a defined volume (1ml) at 37°C for 30min followed by inoculum removal. We prefer this description because with these infection conditions, we find on average well below 100 virus particles that enter the cell (=> This is e.g., reflected in the number of accumulating genomes shown in figure 2A). In contrast, this permits to have enough viruses inside the cell to perform the different "omics" techniques applied in our study to obtain meaningful results at reasonable sequencing depths. This experimental setting was carefully chosen in full awareness of the work by Wang et al., cited by the reviewer, to avoid e.g., overloading the nuclear import rate. Thus, our experimental conditions do not exceed the "MOI of >200" that would affect nuclear import rates. The number (>200) in the Wang et al. study refers to

Full Revision

the number of virus particles inside the cell, the infection condition used in the Wang study was an MOI of 30 bound to Hela cells in the cold for 30min and warmed for 150min which is significantly more virus than we have used in our study. We have expanded the information on the MOI used in the material and methods section to clarify this point (line 685 ff).

8. Figure 5 is of low resolution and was difficult to read.

We thank the reviewer for spotting it. It seems that the Figure quality was compromised during the PDF conversion. We updated the Figures and checked the resolution after PDF conversion.

9. Figure S3 is missing a box from the top set of images indicating the region that is expanded in the detail picture.

We updated Figure S3

10. While I realize it is supplemental data, the difference in quality between the agarose gels shown in Figure S4A and S5A is shocking.

The nature of the experiments is very different and therefore the expected MNase digestion profiles on agarose gels look different. In Figure S5 viral particles were digested with MNase, resulting in a smeary decrease in DNA size. This looks very different from the regular MNase pattern of whole cells that is dominated by the regularly spaced nucleosomes in the heterochromatic regions of the genome. As pVII protects only about 70bp of DNA and its spacing is not as homogenous as the nucleosomal spacing, the pictures shown in Figure S5A were expected as they are.

11. Figure S7 is of low resolution.

We updated the Figures and checked the resolution after PDF conversion.

Reviewer #1 (Significance (Required)):

At least in the field of adenovirus research, this is a very important study. There has been considerable debate in the field regarding the timing and degree of protein VII removal and histone deposition, and the necessity of active transcription for these two events. The data provided in this manuscript clearly shows that some protein VII is removed from early active genes and replaced by nucleosomes, and that these events occur prior to initiation of transcription. The authors speculate that the specific placement of protein VII, a protamine-like protein, on the Ad genome prescribes where nucleosomes are placed. This finding should be of interest to a broad general audience, as it provides novel information on chromatin assembly within mammalian cell. Key words for this reviewer: adenovirus research, HAdV nucleoprotein structure

Reviewer #2 (Evidence, reproducibility and clarity (Required))

The submitted manuscript presents a detailed and comprehensive analysis of the adenoviral nucleoprotein complexes as infection progresses, starting with the "adenosome" assembled with pVII which are then progressively replaced with H3.3.-containing nucleosomes as the infection progresses. The submission presents a combination of in situ and populational analyses of the viral

DNA accessibility and complexes through infection. In brief, the infecting viral genomes are assembled in some 250 adenosomes with pVII, which become progressively replaced as infection progresses with nucleosomes containing H3.3 and acetylated H3K17, starting at the active promoters of the E genes. Chromatin remodeling precedes transcription, and the accessibility differs for genes of different kinetic classes at different times after infection, although there is no correlation between accessibility and H3.3 or acetylation content. Only about 20% of the genomes become transcriptionally active, though, which somewhat complicates the analyses of the populational studies of accessibility and occupancy. Overall, the study is well conceived, performed and presented. A few issues that deserve further analyses and discussion, as described below.

Major issues.

As figure 2 nicely shows, only about 20% of the intranuclear genomes become transcriptionally active. However, MNase and ChIP analyses cannot differentiate these genomes from the 80% that are transcriptionally inactive. The interpretation of the positioning of pVII (figure 4) or the changes in compaction of the adenoviral chromatin at different loci (figure 5) does not appear to consider this heterogeneity other than for a brief comment about the stringent MNase digestion in page 11. The authors favor a model in which the changes in compaction shown in figure 5, at mild MNase digestions, directly correlate with transcription of the respective genes. This could well be correct, and in fact the correlation may be underestimated as 80% of the genomes may not undergo any changes, but it may also be incorrect. The analyses presented cannot differentiate whether the changes in chromatin compaction occur in only a subset of genomes or in all the genomes, regardless of whether they are transcribed or not, or even only in the non-transcribed genomes (which appears extremely unlikely). This intrinsic limitation to the methods used (and I know of no better alternative) should be acknowledged and discussed for the benefit of the reader. This limitation also impacts the analyses of the lack of correlation between H3.3 and acetylated H3K27 occupancy and compaction.

A discussion is amended and located starting from line 571 in the text. "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."

The analysis of the histone ChIP is discussed below.

Perhaps out of necessity to reach the required sensitivity, a high multiplicity of infection was used (although the actual moi is not stated, there are about 25-30 pVII foci/ per nuclei). The presentation, analyses and discussion of the results should emphasize this context. For example, one would presume that at low moi, when only one genome enters each cell, the percentage of transcriptionally active genomes in a given cell will be either 0 or 100%, but the "system" becomes saturated as more and more genomes enter the nucleus at higher moi resulting in only a subset of them being transcriptionally active. Along this line of reasoning, it is intriguing that the percentage of genomes estimated to be in nucleosomes at 4 hpi (14%) approaches the percentage of transcribed genomes.

This issue was also raised by reviewer 1 (see detailed comment above). The reviewer is correct that we chose to use a higher MOI to reach the required sensitivity in our different "Omics" assays. The imaging approach was adapted to reach the morphological equivalence to fit this analysis. We agree that it would be interesting to also study the MOI effect on transcriptional activation (as well as

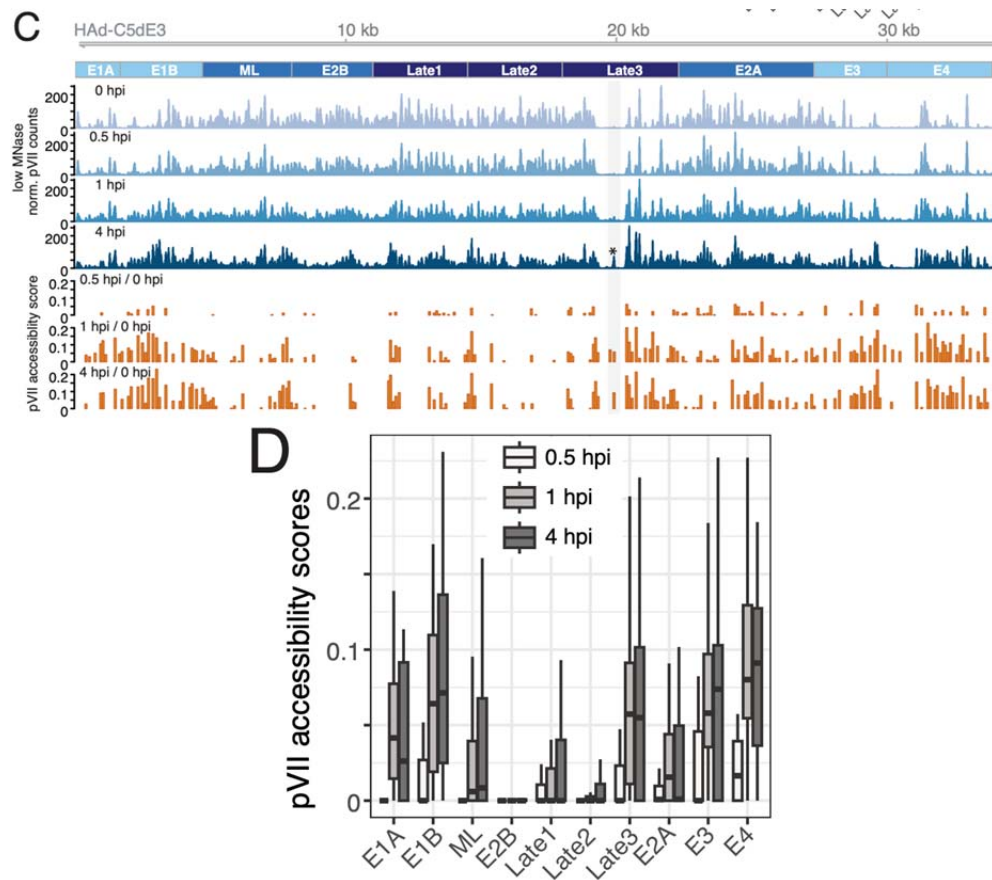
positioning effects, see comment below) but this requires different approaches and will be addressed in a future study. To address this comment (and others in this review) we revised the text in the discussion to emphasize the importance of MOI and possible other effects such as positioning (line 587 ff).

The changes in chromatin compaction presented in figure 5 are in some respect puzzling. The compaction of most of the late genes increases as infection progresses, at least for the first four hours, as the authors discuss. However, the L genes appear to be at least as accessible as the E ones at the early times, when only the E are transcribed to high levels. This appears counterintuitive, and may not be consistent with the main conclusion that increase accessibility to a given gene directly correlates to its transcriptional activity level. The data presented in Figure 5C deserves a more nuanced analysis and discussion, parsing out the changes in accessibility to each given gene at different times from the different accessibility to the different genes at any given time. The later does not appear to support the main conclusion reached by the authors that accessibility to each individual gene correlates with its transcriptional level.

We thank the reviewer for raising this point. While the viral genomes enter the nucleus, the viral chromatin structure is tightly condensed. Therefore, it is unlikely that after nuclear entry the viral chromatin undergoes further compaction. With our analysis, we expect to detect only decompaction of genomic sites relative to 0 hpi, when the virus has not entered the nucleus yet. At some sites and particularly at the Late genes the signal is decreasing, most likely due to normalization to sequencing depth and the variation in the number of viral genomes but not due to changes in compaction. We realized that the negative accessibility scores we used in the study are misleading and give a false impression. Therefore, we changed the analysis in that way, that negative values were not permitted and converted to zeros.

Additionally, we raised the temporal resolution of the analysis and compared the accessibility at all available timepoints against 0 hpi as suggested by the reviewer. Now, we clearly observe, that most accessibility changes are accomplished rapidly after nuclear import, already at 1 hpi and do not change much after, until 4 hpi. Regions of decompaction coincide with early expressed genes and occur before transcription, underscoring the conclusions made in the study. Nevertheless, while most genomic regions covering late genes do not show decompaction, we observed some local sites showing a high accessibility score. As transcription at those sites appears later in the life cycle of the virus, we can only speculate about the function e.g. as enhancer elements.

The Text and Figures were changed accordingly (line 347 ff).

**New legend:**

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.

Minor comments

The authors may wish to highlight in the discussion that the analyses are so far limited to a single adenovirus.

We have taken up the suggestion of the reviewer and included it in the discussion part, starting at line 607:

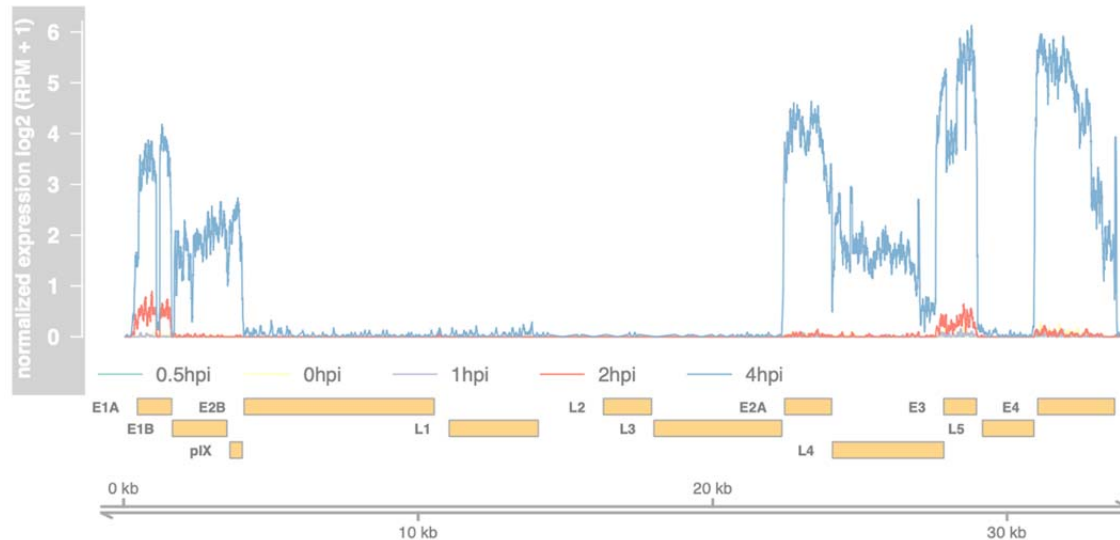
“The structural analysis is still limited to a single adenovirus genotype and it will be interesting to test whether these dynamic changes are conserved among other adenoviruses. Furthermore, reproducing such organization in adenoviral vectors could result in efficient and sustained transgene expression.”

The y-axes in the transcriptome figures (figure 1 B, S2) could be presented in Log(2) scale, such that transcript levels at all times can be appreciated in the same graph (the earlier times are just not visible in a linear scale)

Full Revision

As requested by the reviewer we changed the data to log2 scale. As there is no qualitative difference to the log10 scale, presented in the original version, we would like to keep the figure as it is. To highlight changes at early time points we generated the average expression of early genes in Fig1C.

As an information for the reviewer, we provide here the data plotted as log2 scale.



The (lack of) phenotype of the 24xMS2 binding site recombinant adenovirus used should be shown.

We observed no difference in phenotype between the parental and the MS2 modified virus. We updated Figure S3 and included a gel analysis and specific infectivity data to show this absence of difference.

The kymograph analyses presented in figure 3B appear to show that there are some sites of transcript accumulation sites which do not harbor viral genomes (i.e., green only tracks). Moreover, the interpretation of the TAF1beta-mCherry signal is complicated by the (fully expected) significant "background" signal. Although these results are consistent with those obtained by RNAscope/pVII staining, there appears to be intrinsic limitations to the system, which preclude reaching strong conclusions from it. These confirmatory analyses should probably be moved to the supplementary information section and removed from the main text and figures. The longer evaluation data mentioned as not shown in page 8 is critical to the conclusions and should be shown.

Here we disagree with the reviewer and prefer to keep the data as main figure. All (immobile) transcript accumulation sites are identified by the kymograph analysis and coincide with a genome while free transcripts show a high mobility that is not picked up in the kymograph analysis. This is independently verified in the provided supplemental movies. Depending on the positioning of the genome inside the living cell, accumulating transcripts can appear adjacent to or on top of a genome. This explains the slight shift between RNA and DNA signal for some genomes in the merged image of the kymograph. This is expected as only fully transcribed transcripts and not nascent transcripts are marked by MS2 (the MS2 loops are positioned in the 3'UTR). Also, all genomes (transcribing and non-transcribing) can be identified in the kymograph above background level. To clarify the representation, we have added labels to the kymograph to show which signal is

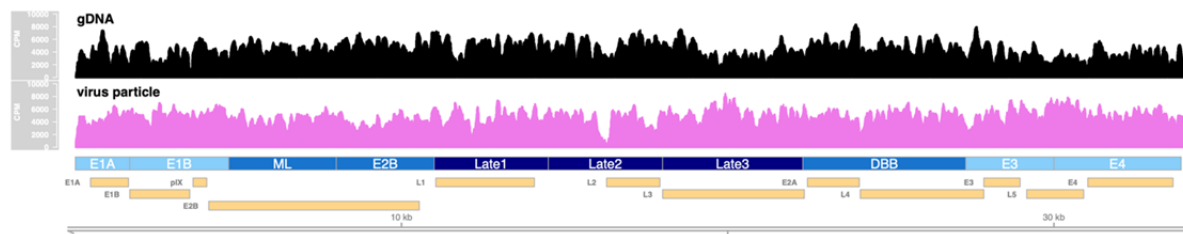
Full Revision

DNA and RNA and a merge respectively. We are convinced that this data set is in strong support of our study, as it is the only technique that permits the discrimination of transcribing and non-transcribing genomes in living cells at real time.

As requested, we have also added two additional examples for a longer observation period (10min) into the supplemental data Fig. S3C.

Although the plot of cleavage frequency presented in figure S5 is clear, it would be beneficial to the reader if the actual peaks were also presented to compare their distribution (if any) in gDNA and virus particle.

In Figure S5 we wanted to test whether the regions lacking pVII peaks are resulting from the absence of pVII, protecting the DNA, and therefore being fully hydrolyzed by MNase, or whether this region is tightly packed by pVII thereby protecting DNA from MNase digestion. To test both possibilities we used a very limited MNase digestion approach, where even free DNA is not fully hydrolyzed, allowing the capture of DNA fragments. Therefore, the sequenced fragments comprise a mixture of protected and un-protected fragments. In this assay, the pVII protected fragments are not fully digested to the monomeric state, but a mix of mono-, di- and other multimers are present. As reflected by the fragment size distribution with the peak between 100-200 bp (Fig S5B), pVII dimers are predominantly enriched when compared to the high MNase digestion used to map pVII positions (compare to Fig4 B). Therefore, the peaks in the S5 data set have a low resolution and do not provide exact pVII positions (see below). Therefore, we would like to keep S5 as it is. We clarified this point in the text (line 279 ff)

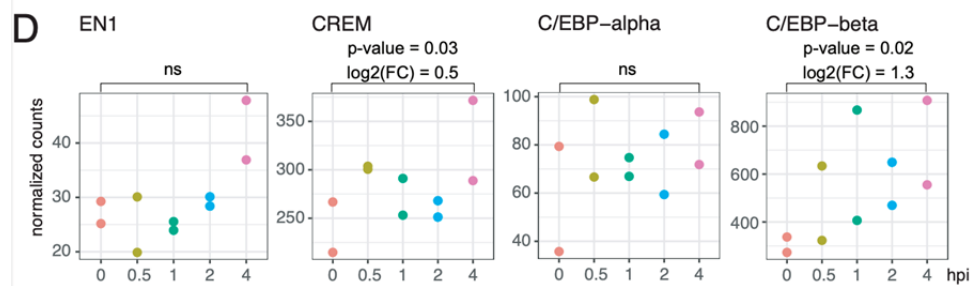


Legend:

Fragment coverage plot of MNase digestions of gDNA (black) or Ad chromatin in virus particles (purple).

The mRNA analyses of selected transcription factors provides little information, as there is no context, there is variability between experiments, and in most cases the changes appear modest. As these results are not critical to the conclusions or analyses, perhaps the authors may wish to remove them from the manuscript. Alternatively, more in-depth analyses would be required.

We agree with the reviewer, that more information for the reader is needed. Therefore, we performed a statistical analysis of expression changes between 0 hpi and 4 hpi of the shown transcription factors using DESeq2. We added the corresponding log2(fold-change) and p-values to the figure. And adapted the text (line 471) and figure accordingly.



Legend:

Gene expression changes of transcription factors over the infection time course. P-values and log₂(fold-changes) from differential gene expression analysis between 4 hpi and 0 hpi using DESeq2 are indicated. ns = not significant

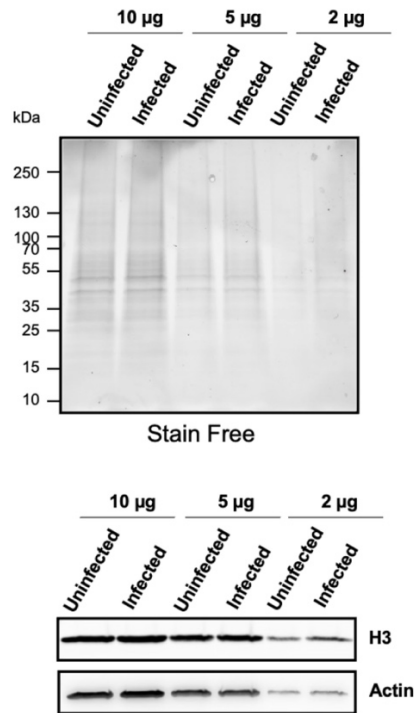
It is unclear why the even distribution of H3.1-flag signal across the genome is considered indicative of no specific recruitment. The results presented are equally consistent with equal incorporation across the genome. Perhaps the authors have some additional information, such as an irrelevant antibody, input DNA, or the like, to support the conclusion. If so, that evidence should be presented and discussed. If not, the interpretation should be revisited. As an added complexity, endogenous H3.1 is normally expressed during S-phase. It is possible that Adenovirus infection may induce higher levels of expression of (untagged)endogenous H3.1, which would outcompete the tagged ectopically expressed histone. These analyses deserve a more nuanced and in-depth analysis.

We have taken several measures in the study to address the concern of the reviewer. We consider timepoint 0 hpi as background control as the viral genome has not entered the nucleus yet. Consistently, we observe very few reads mapped to the Ad genome regardless of antibody and construct used (Fig 6B). Additionally, all samples at 0 hpi cluster together in PCA (Fig 6C) and correlation analysis (Fig S7D)

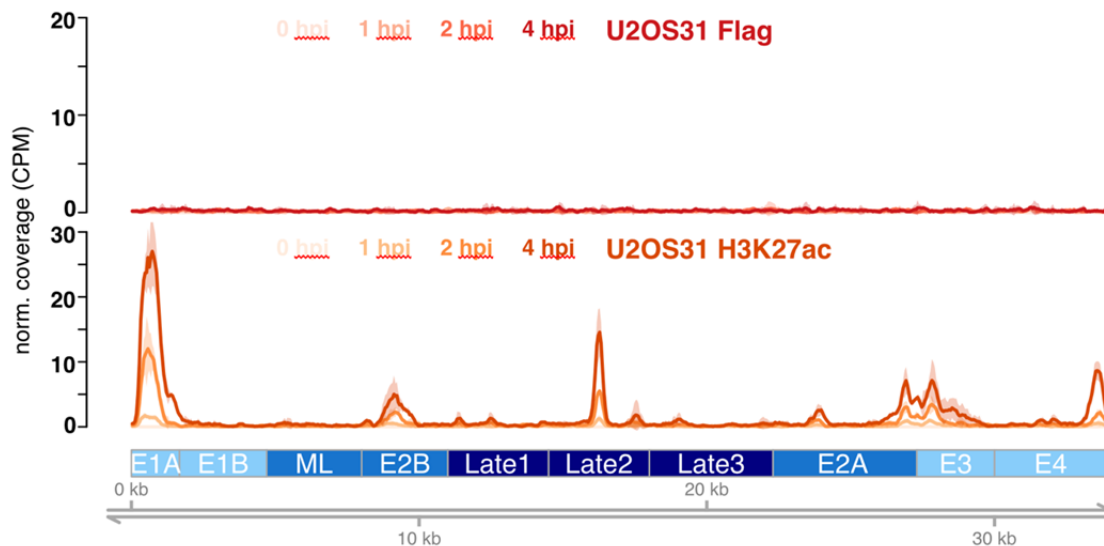
H3.1 Flag tagged samples show at later timepoints (1 - 4hpi) slightly higher percentages of mapped reads to Ad, but plateau already at 1 hpi (Fig 6B) and cluster together in PCA (Fig 6C) and correlation analysis (Fig S7D) with 0 hpi samples. The low background signal starting at 1 hpi for H3.1 might arise due to the change of Ad genome location to the nucleus.

Even though, the number of Ad mapped reads at later timepoints was low in H3.1 Flag tagged samples, it could still be that they accumulate at few sites on the Ad genome indicating a specific deposition. We tested this by plotting the signal across the whole Ad genome (Fig S7E) and zooming into the data (compare scale of H3.3 and H3.1 plot), but we could not detect any reproducible local enrichments. To enable the reader a better comparison between the levels of H3.1 incorporation with H3.3 we put now both on the same scale (Fig 6D and Fig S7D) clearly showing that we cannot detect H3.1 incorporation at Ad genomes in the first 4 hours of infection. The H3.1 signal corresponds to the background noise. We think for two reasons, that it is very unlikely that endogenous H3.1 outcompetes the tagged H3.1:

- 1) The time scale for the cells to transition into S-Phase and upregulate endogenous H3.1 would be only 1-2 hours in our timeseries experiments and therefore too short. To also show these experimentally we amended an experiment for the reviewer that is not included in the manuscript. The Western blots below show that the protein amount of H3 does not increase in the first 4hours of infection. Cells were infected and whole cell extracts were prepared 4hpi.



- 2) As most cells are not in S-phase in our experiments, the expression levels of H3.3 variant is higher than H3.1. With the Flag ChIPs we can clearly show that the tagged H3.3 are not outcompeted by endogenous H3.3. As there is a high sequence similarity between H3.3 and H3.1 it is very unlikely that they behave in that regard differently.



It is highly unlikely that the somewhat higher H3K27ac signal observed in the H3.3 than in the H3.1 expressing cells may result from higher H3.3 occupancy in the viral genome as speculated in page

13. The total levels of H3.3. are unlikely to increase by the ectopically expressed one, and even if they did it is not likely that the occupancy of the viral genome would be limited by the levels of H3.3. This speculation should be removed.

We removed the speculation.

Materials and methods are too concise. A longer more detailed version, as supplementary information, would be highly desirable.

We extended the materials and methods part.

Reviewer #2 (Significance (Required)):

The major strengths of this manuscript lie on its comprehensiveness, using several in situ and populational approaches to address biologically critical questions regarding the regulation of viral replication by chromatin and epigenetics. Experiments appears very well designed and performed and are mostly clearly presented. The interpretation analyses and discussion of the results may benefit from a more nuanced analysis of the issues posed by the existence of different populations of viral genomes in the cells infected at high moi and the accessibility across different genes at any given time versus the levels of transcription of the different genes, which appears not to be fully consistent with one of the main conclusions reached.

This study makes a very significant contribution, describing the dynamic changes in the adenoviral nucleoprotein complexes at the early times of infection and providing a full description of both the adenosomes and the nucleosomes in more and less transcribed loci. The results are properly analyzed in context of what is known about the regulation of viral gene transcription by chromatin dynamics in other systems, including similarities and differences. This study is likely to be of high interest to a wide audience, ranging from virologist to epigeneticists, to those working in gene therapy and vectored vaccines.

Reviewer #3 (Evidence, reproducibility and clarity (Required)):

The manuscript "Adenoviral chromatin organization primes for early gene activation" combines RNA-seq, MNase-seq, ChIP-seq, and single genome and transcript imaging (immunofluorescence, RNA-scope, and live cell techniques) during early Adenovirus infection in vitro to characterise the spatiotemporal dynamics of viral chromatin organisation and association with gene transcription. The manuscript is an interesting read and the authors have combined multiple complimentary techniques to make a substantial contribution to understanding the early events occurring after nuclear import of viral genomes.

Adenoviruses are important causes of human and animal pathology, are a useful model of non-integrating extra-chromosomal DNA virus infection in mammalian cells, and are useful vectors for vaccination and the discoveries may influence gene therapy DNA vector design. The chromatin organisation in adenovirus infection is distinct from other DNA viruses, and is relatively poorly understood compared to, for example, SV40 or herpesviruses. The manuscript describes an early transition from purely viral chromatin with Adenovirus protein pVII packaging the virus in virions, to a viral-human hybrid chromatin pattern with apparently strategically positioned H3.3 nucleosomes and viral pVII "Adenosomes" in the early hours after nuclear import of the viral genome. The data shows

that packaged Adenoviruses are in a transcriptionally accessible form and gene expression occurs rapidly after infection, the combination of the MNase-seq data with ChIP-seq data is particularly interesting demonstrating an average ~238 adenosomes positioned by specific DNA code protecting 60-70bp of DNA, and that the genome is accessible at loci that also decondense on infection, with adenosomes being replaced by cellular H3.3 containing nucleosomes at distinct sites. Particularly they show that +1 H3K27 acetylated nucleosomes are acquired at the TSS of key early genes.

The authors argue that their spatiotemporal data imply that this chromatin transition "primes" for early gene transcription. The manuscript is well written, uncovers important viral chromatin biology by combining multiple experimental techniques, and the data is generally very clearly presented. A few comments follow.

Major concerns:

- Abstract and Title:

- o the abstract and title suggest that because the chromatin changes are observed coincidentally or before transcriptional changes, and that this means that these chromatin changes "prime" (title) or are "required" and play a "central role" (abstract) in early gene expression. The temporal relationship would be consistent with chromatin changes being required for transcriptional changes, but do not imply necessity. Experiments to demonstrate the necessity of these changes for early gene transcription are lacking, and I recommend amending the text or additional experiments to provide this evidence directly.

We observe a clear timing of events, with chromatin opening, nucleosome assembly at the 5' end of the gene followed by transcriptional activation, suggesting that these structural changes are essential for gene activation. Still, we cannot prove the direct dependency. Therefore we toned down the title of our manuscript and formulate the findings more conservatively.

The title now reads: "Changes in adenoviral chromatin organization precede early gene activation"

- Results:

- o The IF data in Fig S1 is convincing, showing viral particles are accessible quickly in the nucleus. Although no statistics are provided for S1B and C, pVII foci appear at 0.5hpi and appear to mostly accumulate between 0.5hpi and 1hpi with further import between 1hpi and 4hpi. Can the authors be sure that a single pVII IF focus represents a single genome? If genomes tend to aggregate as they accumulate the number of foci per nucleus may not increase linearly with the number of genomes imported. Have the authors considered analysing the intensity of the individual pVII foci over the time points? A related question is whether the authors assume that all packaged virions contain intact complete viral genomes? Many viruses comprise some mixture of complete and incomplete packaged genomes, and the subsequent analyses determine the proportion of transcriptionally active copies with RNA-Scope to a single transcript E1A which lies at one end of the viral genome. Please comment explicitly on whether this is assumed and whether this assumption is realistic in light of known Adenovirus biology.

We appreciate the reviewer's concern. Several studies in the adenovirus field have shown equivalence between protein VII nuclear foci and individual genomes, including our own (PMID: 26332038). Probably the most accurate study was performed by Daniel Engels lab PMID: 19406166, who used nuclear protein VII foci to titrate viral as well as vector genomes. In contrast, a different study from Patrick Hearings lab PMID: 21345950 showed that past 4hpi, the number of nuclear protein VII foci gradually declines. Based on our experience and because our study is limited to 4 hpi we are confident that protein VII foci accurately reflect individual viral genomes.

Concerning genome packaging, adenovirus particles contain a single viral genome that is protected at each end by a covalently attached protein preventing its degradation. The packaging of adenoviruses is extremely efficient and only complete genomes are packaged into fully assembled particles. All viruses used in this study have been purified by double CsCl gradient purification. This density gradient based purification protocol removes all particles that are either empty or damaged or would contain partial genomes.

o The RNA-Seq data in Fig 1 and Fig S2 and Table S1 demonstrates transcription of early genes is barely observable at 1hpi but is observable by 2hpi and is clearly much increased by 4hpi. Fig 2C, visualising pVII foci directly within single cells, suggests that approximately 80% of foci are observed by 1hpi and a further 20% between 1hpi and 2hpi and little thereafter. These data convincingly demonstrate that nuclear import is rapid, typically occurring in the first hour. The E1A RNA-Scope data in figure 2, visualising individual mRNA transcripts of E1A, is more sensitive than the bulk RNA-Seq, and shows transcripts at 1hpi with clearly discernible transcription by 2hpi (2A&D) which suggests that transcription occurs early, by 2hpi. Thus transcription lags nuclear genome import by approximately one hour by these methods. However, the conclusions of the subsequent analyses depend on the chromatin changes clearly preceding, rather than being approximately coincident with transcription, therefore transcription being evident by 2hpi is relevant as figure 6A and D suggest that the chromatin remodelling is subtle before 2hpi on the bulk sequencing analyses. The authors should comment on this given the importance to their argument.

As stated by the reviewer we observe a clear lag between nuclear import and transcriptional activation. And we do also observe the largest changes in nucleosome occupancy (ChIP-seq data) between 1 and 2 hpi (Fig6A and D). Compared to 0hpi, we observe the strongest increase of nucleosome occupancy between 1hpi and 2hpi (4-8fold effect), whereas depending on the area a 2-3fold increase in occupancy can be observed from 2hpi to 4hpi (Fig6D). An effect that one would expect with chromatin structure preceding gene activation. Furthermore, the timing of nucleosome assembly perfectly matches the increase of MNase accessibility at 1 hpi, supporting our conclusions.

o The validation of the E1A probe specificity in Fig 2B looks convincing, but there are no data presented for multiple cells to reassure that this image is representative. The equivalent figure for 2D for the Ad5-GFP control would address this.

We include a large field overview with multiple cells for virus and vector control as new supplemental figure S2B showing that the RNAscope detection of the E1A transcript is highly specific.

o Figure 2E is presented as a colocalization analysis but appears to be a ratio of mRNA foci to pVII foci per cell. If this is an incorrect interpretation then some clarification in the figure legend would be helpful. If this interpretation of these data is correct, then it is not truly a colocalization analysis, as a single genome may give rise to multiple transcripts and so a ratio <1 could be expected even if all pVII foci were transcribing genomes. In addition, the text suggests that because there is no statistically significant difference between 2hpi and 4hpi a plateau has been reached. However, the difference between 1h and 2h is barely significant ($p=0.04$) and the mean is increased between 2h and 4h, albeit non-significantly, so the plateau is not convincingly demonstrated. If the authors wish to perform a colocalization analysis rather than a ratio, they might assign each transcript to the nearest pVII IF focus within the nucleus and count the proportion of pVII foci with any transcripts assigned over time. Alternatively, they can amend the description of this analysis in the figure and text.

We apologize that this figure was not clear. The data are based on real colocalizations and represent the number of pVII dots positive for E1A normalized with the total number of nuclear pVII. We have clarified the figure legend accordingly.

o The live cell imaging experiments are elegant and convincing, but the agreement in Fig 3D of the % colocalization in MS2-BP data with the RNA-scope data is potentially misleading for the reasons outlined in the prior comment. Is the data in Fig 2E the same as the data in the right hand panel of Fig 3D. If so please comment on the n discrepancy (n=30 in 2E vs n=22 in 3D). The observation that 20% of genomes are transcriptionally active, via bursting or otherwise, is interesting, and would be consistent with the Suomalainen et al reference. The authors discuss two hypotheses to explain these findings: transcriptional bursting or a subset ~20% of genomes being transcriptionally active. This is an interesting and begs the question as to why this may occur. Assuming all imported genomes are intact (previous comment), it appears from the presented images that the foci at the radial periphery of the nucleus may be more frequently transcriptionally active, despite the nuclear periphery being enriched for heterochromatin. The authors might consider analysing the radial position of their TAF1B-mCherry genomes (active and inactive) as this might support position effect variegation rather than bursting as an explanation and they appear to already have the data to perform such analyses.

o In the presented images (Fig 3A and Fig S3) it appears a higher proportion of genomes than 20% appear to be transcriptionally active, particularly in the low MOI experiment. The authors may wish to comment on this and quantify whether the proportion of transcribing genomes was affected by the input MOI.

This and the previous comment concerning the influence of MOI, transcriptional bursting and the positioning effect of the genome on the transcriptional activity have also been in part raised above. As stated in our response to reviewer 1 we have used a high MOI in our experiments to have equivalence between all experimental approaches. We agree with the reviewers that all aspects (dose, bursting and positioning) merit a detailed investigation, which we plan in future studies. To be consistent and comparable in our comprehensive approach we decided to not include such studies here as they would address a different question. Nevertheless, to address this (and the above) comments we now mention positioning effects in the results (line 214) and enlarged the discussion (line 587 ff) where we especially raised awareness that such pertinent questions can be addressed with the tools presented in our study.

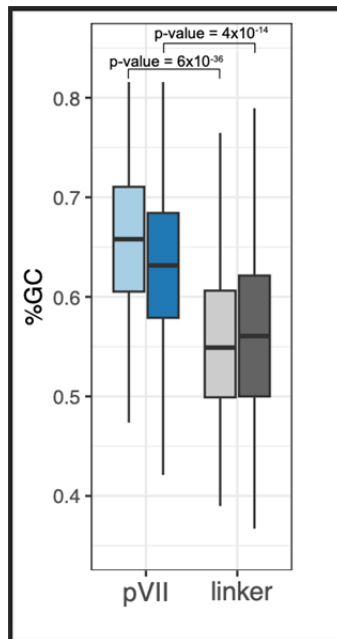
We also decided to visually separate the comparison of MS2 and RNAscope data to avoid misleading the reader. Furthermore, the RNAscope data have been replaced. The RNAscope data are indeed from Fig. 2. The difference in n was due to our mistake showing two different normalized data sets. Data were either normalized using total amount of nuclear protein VII (Fig. 2E) or the total amount of nuclear E1A signals (Fig 3D), which due to the more heterogenous signal did not include all cells. In the updated version both figures display data normalized by total amount of nuclear protein VII

o Fig 4C suggests that there is a large GC preference (or bias) in the pVII occupied regions. The authors may wish to comment on this and present a track with Adenovirus GC composition in Fig 4D.

We thank the reviewer for raising this point. As suggested by the reviewer we analysed the GC content under pVII peaks and in the linker DNA. Indeed, pVII occupied regions have a significant

Full Revision

higher GC content indicating that pVII preferentially positions at GC rich regions. We included this analysis as an additional Figure 4E (line 302 f).



Legend:

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.

o Figure 6 presents convincing data showing H3.3. nucleosome positioning and acetylation at E1A and the data is nicely presented showing these changes occur early being observable by 2 and 4 hpi. Again, these changes are not convincingly prior to early gene activation but are certainly occurring early, and may occur prior to early gene activation at the level of individual foci, however, this is not demonstrated definitively.

This question belongs to the same context addressed by the reviewer above. Please refer to the answer given above.

Minor comments:

- Introduction:

- o Paragraph 1 - Introduction for DNA viruses in general, but the authors appear to be talking about Adenoviruses specifically, "little is known about the structural organization of the genome" and "nuclear viral genomes could undergo different parallel fates", arguably these statements are not accurate for other DNA viruses (e.g. Epstein Barr Virus) suggest amending the wording for clarity.

The manuscript text was updated as suggested.

- o paragraph 2 - Why do the authors say that Adenoviruses are prototypic DNA viruses?

We removed the term prototypic.

Full Revision

- o Paragraph 3 - A recent study is referenced but multiple references are given.

The references were updated

- o "Protein VII stays associated with the viral genome imported to the nucleus, while pV dissociates from the viral DNA following ubiquitylation (Puntener et al., 2011). The fate of the μ -peptide is not known". - The reference suggests that pV dissociates on entry to the cytoplasm and during capsid disassembly at the nuclear pore. I find this sentence confusing as it doesn't make it clear that pV is lost before nuclear entry which is important for interpreting the data.

We clarified this in the manuscript text

- Results:
 - o Figure 5 is almost unreadable due to low resolution.

We updated the Figures and checked the resolution after PDF conversion.

- o Reference to Fig 4C in text comes after Fig4D.

The order of Figure panels was changed accordingly.

Reviewer #3 (Significance (Required)):

The manuscript is well written, uncovers important viral chromatin biology by combining multiple experimental techniques, and the data is generally very clearly presented

Dear Dr. Längst,

Thank you for submitting your revised Review Commons manuscript to The EMBO Journal. Your manuscript has now been seen by all original reviewers, who find that their main concerns have been addressed and now recommend publication of the manuscript. I will therefore be happy to accept the manuscript for publication in The EMBO Journal after addressing the minor comments by reviewer #2 and reformatting of the manuscript along the guidelines included in the attached document.

Please feel free to contact me if you have any further questions regarding this final editorial revision. You can use the link below to upload the revised files.

Thank you for the opportunity to consider your work for publication, and I look forward to receiving your revised manuscript.

With best regards,

Ieva

Ieva Gailite, PhD
Senior Scientific Editor
The EMBO Journal
Meyerhofstrasse 1
D-69117 Heidelberg
Tel: +4962218891309
i.gailite@embojournal.org

Instructions for preparing your revised manuscript:

Please make sure you upload a letter of response to the referees' comments together with the revised manuscript.

Please also check that the title and abstract of the manuscript are brief, yet explicit, even to non-specialists.

When assembling figures, please refer to our figure preparation guideline in order to ensure proper formatting and readability in print as well as on screen:

<https://bit.ly/EMBOPressFigurePreparationGuideline>

See also figure legend guidelines: <https://www.embopress.org/page/journal/14602075/authorguide#figureformat>

IMPORTANT: When you send the revision we will require

- a point-by-point response to the referees' comments, with a detailed description of the changes made (as a word file).
- a word file of the manuscript text.
- individual production quality figure files (one file per figure)
- a complete author checklist, which you can download from our author guidelines (<https://www.embopress.org/page/journal/14602075/authorguide>).
- Expanded View files (replacing Supplementary Information)

Please see out instructions to authors

<https://www.embopress.org/page/journal/14602075/authorguide#expandedview>

Please remember: Digital image enhancement is acceptable practice, as long as it accurately represents the original data and conforms to community standards. If a figure has been subjected to significant electronic manipulation, this must be noted in the figure legend or in the 'Materials and Methods' section. The editors reserve the right to request original versions of figures and the original images that were used to assemble the figure.

Further information is available in our Guide For Authors: <https://www.embopress.org/page/journal/14602075/authorguide>

Revision to The EMBO Journal should be submitted online within 90 days, unless an extension has been requested and approved by the editor; please click on the link below to submit the revision online before 20th Jul 2023:

Link Not Available

Referee #1:

The authors have submitted a revised manuscript addressing the nucleoprotein structure of human adenovirus during the very early stages of infection, and its relationship to onset of expression of viral genes. They show that in the virion and the newly-infecting DNA, protein VII is precisely positioned at G-C rich sites on the viral DNA, with greater accessibility at early gene promoters compared to other regions. Nucleosomes containing H3.3 replace protein VII near the transcription start sites of early genes, which are then acetylated. Association with histones and nucleosomes occurs appear to precede transcription. These studies confirm and greatly expand on results already in the literature.

The authors have done an admirable job of addressing the reviewer's concerns, through addition of some new data, or re-analysis or re-explanation of previously-presented data. The text of the manuscript has been revised to improve clarity. I have no additional comments or concerns.

Referee #2:

The revised submission has satisfactorily addressed all the previous critiques. The additional or edited analyses and text add to the quality of an already excellent manuscript. This reviewer has only minor suggestions to the authors.

The final sentence in Ad genome accessibility in materials and methods appears to contain a typo - it probable was meant to read "further compaction..." , not "decompaction". This reviewer is not keen on the change to the accessibility scores removing negative values, as it is based on a presumption that further compaction is not possible, assumption which is not tested. The approach used is clearly described and the reader can evaluate it and agree or not with it, so it is up to the authors to decide whether to revert to the original analyses or keep the current version.

From the title and through the text, the changes in adenoviral chromatin organization are very consistent with regulation of transcription competency, not actual transcription. Perhaps the use of the specific wording could help the reader.

When comparing the differences in the changes in chromatin accessibility between the previously reported HSV-1 and the current work on adenovirus, the authors may wish to consider that adenoviruses replicate as individual genomes, whereas herpesviruses create long concatemeric ones, and that PTM are also distributed through entire HSV-1 genomes, not localized to specific loci as shown now for adenovirus.

Referee #3:

The authors have addressed all of our comments satisfactorily.

Rev_Com_number: RC-2023-01819

New_manu_number: EMBOJ-2023-114162

Corr_author: Längst

Title: Changes in adenoviral chromatin organization precede early gene activation

Manuscript EMBOJ-2023-114162

point-by-point response to the referees' comments

Referee #1:

The authors have submitted a revised manuscript addressing the nucleoprotein structure of human adenovirus during the very early stages of infection, and its relationship to onset of expression of viral genes. They show that in the virion and the newly-infecting DNA, protein VII is precisely positioned at G-C rich sites on the viral DNA, with greater accessibility at early gene promoters compared to other regions. Nucleosomes containing H3.3 replace protein VII near the transcription start sites of early genes, which are then acetylated. Association with histones and nucleosomes occurs and appears to precede transcription. These studies confirm and greatly expand on results already in the literature.

The authors have done an admirable job of addressing the reviewer's concerns, through addition of some new data, or re-analysis or re-explanation of previously-presented data. The text of the manuscript has been revised to improve clarity. I have no additional comments or concerns.

[Author comment: We thank the reviewer for reviewing the manuscript and for the positive input.](#)

Referee #2:

The revised submission has satisfactorily addressed all the previous critiques. The additional or edited analyses and text add to the quality of an already excellent manuscript. This reviewer has only minor suggestions to the authors.

The final sentence in Ad genome accessibility in materials and methods appears to contain a typo - it probably was meant to read "further compaction..." , not "decompaction". This reviewer is not keen on the change to the accessibility scores removing negative values, as it is based on a presumption that further compaction is not possible, assumption which is not tested. The approach used is clearly described and the reader can evaluate it and agree or not with it, so it is up to the authors to decide whether to revert to the original analyses or keep the current version.

[Author comment: We corrected the typo and kept the figure without the negative values as the final figure version for the manuscript.](#)

From the title and through the text, the changes in adenoviral chromatin organization are very consistent with regulation of transcription competency, not actual transcription. Perhaps the use of the specific wording could help the reader.

[Author comment: We agree with the reviewer and corrected the wording in the text, when necessary for comprehension.](#)

When comparing the differences in the changes in chromatin accessibility between the previously reported HSV-1 and the current work on adenovirus, the authors may wish to consider that adenoviruses replicate as individual genomes, whereas herpesviruses create long concatemeric ones, and that PTM are also distributed through entire HSV-1 genomes, not localized to specific loci as shown now for adenovirus.

Author comment: As we are studying the early time points of infection, before replication, we did not consider including the information about HSV-1. We thank the reviewer for reviewing the manuscript and for the helpful suggestions.

Referee #3:

The authors have addressed all of our comments satisfactorily.

Author comment: We thank the reviewers for taking the time and evaluating and improving our manuscript.

With best regards
Gernot Längst and Harald Wodrich

Dear Dr. Längst,

Thank you for addressing most of the final editorial points. I am now pleased to inform you that your manuscript has been accepted for publication.

I will look into the synopsis text in the next couple of days and let you know if any edits to the journal style are needed.

Please note that it is EMBO Journal policy for the transcript of the editorial process (containing referee reports and your response letter) to be published as an online supplement to each paper. If you do NOT want this, you will need to inform the Editorial Office via email immediately. More information is available here:
https://emboj.embopress.org/about#Transparent_Process

Your manuscript will be processed for publication in the journal by EMBO Press. Manuscripts in the PDF and electronic editions of The EMBO Journal will be copy edited, and you will be provided with page proofs prior to publication. Please note that supplementary information is not included in the proofs.

Please note that you will be contacted by Wiley Author Services to complete licensing and payment information. The 'Page Charges Authorization Form' is available here: https://www.embopress.org/pb-assets/embo-site/tej_apc.pdf

EMBO Press participates in many Publish and Read agreements that allow authors to publish Open Access with reduced/no publication charges. Check your eligibility: <https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/affiliation-policies-payments/index.html>

Should you be planning a Press Release on your article, please get in contact with embojournal@wiley.com as early as possible, in order to coordinate publication and release dates.

If you have any questions, please do not hesitate to call or email the Editorial Office. Thank you for this contribution to The EMBO Journal and congratulations on a successful publication!

Best regards,

Ieva

Ieva Gailite, PhD
Senior Scientific Editor
The EMBO Journal
Meyerhofstrasse 1
D-69117 Heidelberg
Tel: +4962218891309
i.gailite@embojournal.org

Rev_Com_number: RC-2023-01819
New_manu_number: EMBOJ-2023-114162R
Corr_author: Längst
Title: Changes in adenoviral chromatin organization precede early gene activation

EMBO Press Author Checklist

Corresponding Author Name: Gernot Längst
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2023-114162

USEFUL LINKS FOR COMPLETING THIS FORM

[The EMBO Journal - Author Guidelines](#)
[EMBO Reports - Author Guidelines](#)
[Molecular Systems Biology - Author Guidelines](#)
[EMBO Molecular Medicine - Author Guidelines](#)

Reporting Checklist for Life Science Articles (updated January)

This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: [10.31222/osf.io/9sm4x](https://doi.org/10.31222/osf.io/9sm4x)). Please follow the journal's guidelines in preparing your manuscript. **Please note that a copy of this checklist will be published alongside your article.**

Abridged guidelines for figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if $n < 5$, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements.
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
 - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
 - are tests one-sided or two-sided?
 - are there adjustments for multiple comparisons?
 - exact statistical test results, e.g., P values = x but not P values $< x$;
 - definition of 'center values' as median or average;
 - definition of error bars as s.d. or s.e.m.

Please complete ALL of the questions below.

Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Yes	Section "Data Availability"
Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For antibodies provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and/or clone number - Non-commercial: RRID or citation	Yes	Materials and Methods
DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Not Applicable	
Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository OR supplier name, catalog number, clone number, and/OR RRID.	Yes	Materials and Methods
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Not Applicable	
Report if the cell lines were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Not Applicable	
Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository OR supplier name, catalog number, clone number, OR RRID.	Not Applicable	
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail housing and husbandry conditions .	Not Applicable	
Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Not Applicable	
Microbes: provide species and strain, unique accession number if available, and source.	Yes	Virus Information in Materials and Methods
Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Not Applicable	
Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgments section?	Yes	Acknowledgements

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been pre-registered , provide DOI in the manuscript . For clinical trials, provide the trial registration number OR cite DOI.	Not Applicable	
Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	Not Applicable	
Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if external detailed step-by-step protocols are available.	Not Applicable	
Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about sample size estimate even if no statistical methods were used.	Yes	Materials and Methods and Result sections
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, have they been described?	Not Applicable	
Include a statement about blinding even if no blinding was done.	Not Applicable	
Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Not Applicable	
If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.		
For every figure, are statistical tests justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Materials and Methods and Result sections
Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was replicated in laboratory.	Yes	Figure Legends
In the figure legends: define whether data describe technical or biological replicates .	Yes	Figure Legends

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving human participants : State details of authority granting ethics approval (IRB or equivalent committee(s), provide reference number for approval.	Not Applicable	
Studies involving human participants : Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Not Applicable	
Studies involving human participants : For publication of patient photos , include a statement confirming that consent to publish was obtained.	Not Applicable	
Studies involving experimental animals : State details of authority granting ethics approval (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Not Applicable	
Studies involving specimen and field samples : State if relevant permits obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	
Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of select agents and toxins (CDC): https://www.selectagents.gov/sat/list.htm .	Not Applicable	
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	
If a study is subject to dual use research of concern regulations, is the name of the authority granting approval and reference number for the regulatory approval provided in the manuscript?	Not Applicable	

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., ICMJE, MIBBI, ARRIVE, PRISMA) have been followed or provided.	Yes	
For tumor marker prognostic studies , we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Not Applicable	
For phase II and III randomized controlled trials , please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	Not Applicable	

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have primary datasets been deposited according to the journal's guidelines (see Data Deposition section) and the respective accession numbers provided in the Data Availability Section?	Yes	Data Availability
Were human clinical and genomic datasets deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Not Applicable	
Are computational models that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Yes	
If publicly available data were reused, provide the respective data citations in the reference list .	Not Applicable	